

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:44:16 ; Search time 88 seconds
(without alignments)
505.039 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSEPDCAKNESLYGRKYI.....QGFVPGDQVLCQGSQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1635	100.0	280	21	AA1984378
2	862	52.7	284	23	ABP64881
3	716	43.8	245	22	ABP6074
4	716	43.8	245	22	AA195383
5	716	43.8	245	22	AAU20437
6	622	38.0	245	22	ABP64785
7	478.5	29.3	188	22	AA13792
8	413.5	25.3	328	22	ABP66282
9	349	21.3	116	22	ABG11960

10	328	20.1	746	22	ABB67068
11	328	20.1	765	22	ABB64456
12	327	20.0	497	22	ABB67109
13	325.5	19.9	177	22	ABP65964
14	314	19.2	292	20	AAW88807
15	314	19.2	292	22	ABB50715
16	313	19.1	334	22	ABB50720
17	313	19.1	386	20	AA131985
18	313	19.1	386	24	ABU07726
19	313	19.1	417	23	ABP41908
20	311.5	19.1	444	23	ABB57313
21	311.5	19.1	444	23	ABB57368
22	306.5	18.7	314	17	AAW06724
23	306.5	18.7	314	22	AAE03745
24	306.5	18.7	314	22	AAE03745
25	306.5	18.7	314	23	ABG31150
26	306.5	18.7	314	23	ABG31150
27	305.5	18.7	316	24	ABU71010
28	305.5	18.7	461	22	AAW78958
29	305.5	18.7	478	20	AA173975
30	305.5	18.7	486	22	ABP11886
31	305.5	18.7	486	22	AAW79942
32	305	18.7	495	22	ABB57748
33	305	18.7	495	22	ABB57752
34	304.5	18.6	591	19	AAW41418
35	304.5	18.6	605	19	AAW41419
36	304	18.6	444	21	AAW78419
37	303.5	18.6	633	24	ABU11586
38	303.5	18.6	738	24	ABU11765
39	301	18.4	266	23	ABG70071
40	301	18.4	596	23	ABG61825
41	297.5	18.2	336	24	AAQ16355
42	297.5	18.2	341	22	ABP93978
43	297.5	18.2	341	22	ABP93978
44	297.5	18.2	341	22	ABP66598
45	293	17.9	454	17	AAR88904

ALIGNMENTS

RESULT 1

AA1984378

ID AA1984378 standard; Protein; 280 AA.

XX AC

XX AC

XX AC

DT 12-JUL-2000 (first entry)

XX Amino acid sequence of a human LIM domain protein homologue.

XX Human; LIM domain protein homologue; LIM domain; LIM protein; I protein;

XX M protein; Incyte clone 1925017; cancer; reproductive disorder;

XX autoimmune disorder.

XX Homo sapiens.

XX Key

PH Modified-site

FT Location/Qualifiers

FT 11

FT /note= "potential N-glycosylation site"

FT 15

FT /note= "potential tyrosine kinase phosphorylation site"

FT 39

FT /note= "potential casein kinase II phosphorylation site"

FT 40..98

FT /note= "LIM domain"

FT 74

FT /note= "potential casein kinase II phosphorylation site"

FT 81

FT /note= "potential casein kinase II phosphorylation site"

FT 101..159

FT /note= "LIM domain"

FT 112

FT Modified-site /note= "potential protein kinase C phosphorylation site"
 FT 121 /note= "potential casein kinase II phosphorylation site"
 FT Modified-site 130 /note= "potential casein kinase II phosphorylation site"
 FT Domain 162..218 /note= "LIM domain"
 FT Modified-site 175 /note= "potential casein kinase II or protien kinase C phosphorylation site"
 FT Modified-site 201 /note= "potential casein kinase II phosphorylation site"
 FT Modified-site 202 /note= "potential casein kinase II phosphorylation site"
 FT Domain 221..275 /note= "LIM domain"
 FT Modified-site 223 /note= "potential protein kinase C phosphorylation site"
 FT Modified-site 238 /note= "potential casein kinase II phosphorylation site"
 XX WO200014231-A1.
 PN 16-MAR-2000.
 XX 01-SEP-1999; 99WO-US20239.
 XX 03-SEP-1998; 98US-0155260.
 XX (INCY-) INCYTE PHARM INC.
 XX Tang YT, Guegler KJ, Corley NC, Patterson C;
 XX WPI; 2000-256982/22.
 DR N-PSDB; AAZ99798.
 XX A purified polypeptide designated LDPH (LIM domain protein homolog),
 PT useful for treating, preventing, and diagnosing various cancers,
 PT reproductive disorders and autoimmune disorders.
 XX Claim 1; Page 52-53; 59pp; English.
 XX The present sequence represents a human LIM domain protein homologue.
 CC The LIM domain is so named because it was first described in three
 CC proteins from Drosophila melanogaster, designated L, I, and M. The LIM
 CC domain is rich in cysteine. Nucleic acids encoding the LIM protein
 CC homologue were first identified in Incyte clone 1925017 from a breast
 CC tissue cDNA library. The LIM polynucleotide and polypeptide are
 CC useful for treating, preventing, and diagnosing various cancers,
 CC reproductive disorders and autoimmune disorders.
 XX Sequence 280 AA;
 Query Match 100.0%; Score 1635; DB 21; Length 280;
 Best Local Similarity 100.0%; Pred.No. 1.4e-127;
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MSESFDCAKNEISLYGRKIYQTDSPGVCPCYDNTFANTCAEQQLIGHDSRELFYEDRH 60
 Db 1 MSESFDCAKNEISLYGRKIYQTDSPGVCPCYDNTFANTCAEQQLIGHDSRELFYEDRH 60
 Qy 61 FHEGCFRCRCQRLADEPFTQDSPELLCNDVCYSAFSSQCSAGETVMPGSRKLEYGGQ 120
 Db 61 FHEGCFRCRCQRLADEPFTQDSPELLCNDVCYSAFSSQCSAGETVMPGSRKLEYGGQ 120
 Qy 121 TWHEHCFLCSCGEPLGSRSPVDPKGAHYCVPCYENKFAPRCARCSKTLTQGGVYTRDQP 180
 Db 121 TWHEHCFLCSCGEPLGSRSPVDPKGAHYCVPCYENKFAPRCARCSKTLTQGGVYTRDQP 180
 Qy 181 WHRECLVTCGQTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPVGLGGKYYVSFE 240
 Db 181 WHRECLVTCGQTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPVGLGGKYYVSFE 240

Qy 241 DRHHHNCFCARGSTSLVGQGFVPDGDQVLCQCSQAGP 280
 Db 241 DRHHHNCFCARGSTSLVGQGFVPDGDQVLCQCSQAGP 280
 RESULT 2
 ABP64881
 ID ABP64881 standard; Protein; 284 AA.
 AC ABP64881;
 DT 25-FEB-2003 (first entry)
 XX Human protein SEQ ID 541.
 DE Human; expressed sequence tag; EST;
 KW haematopoietic disorder; central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;
 KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
 KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective.
 XX Homo sapiens.
 OS WO200259260-A2.
 PN 01-AUG-2002.
 PD 16-NOV-2001; 2001WO-US42950.
 PF 17-NOV-2000; 2000US-0714936.
 PR (HYSE-) HYSEQ INC.
 XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX WPI; 2002-590824/63.
 DR N-PSDB; ABQ99467.
 XX New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.
 XX Claim 20; SEQ ID 541; 394pp; English.
 CC The present invention relates to novel human coding sequences
 CC (ABQ99268-ABQ99608) and proteins (ABP64682-ABP65022). The sequences are
 CC useful in therapeutic, diagnostic and research methods. The
 CC polynucleotides may be used in the field of molecular biology as
 CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
 CC for the recombinant production of protein, or in generation of anti-sense
 CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
 CC sequence tags (ESTs) for identifying expressed genes or for physical
 CC mapping of the human genome. The proteins may be used as molecular weight
 CC markers, or as nutritional sources or supplements. The proteins may be
 CC used to maintain and expand cell population in a totipotential or
 CC pluripotential state useful for re-engineering damaged or diseased
 CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
 CC development of bio-sensors. The polynucleotides and proteins are useful
 CC for preventing, treating or ameliorating disorders involving aberrant
 CC protein expression or biological activity, e.g. haematopoietic disorders,
 CC central/peripheral nervous system diseases, mechanical and traumatic
 CC disorders, non-healing wounds, immune deficiencies and disorders,
 CC infectious diseases caused by viral, bacterial or fungal infection,
 CC autoimmune disorders, allergic reactions and conditions, coagulation
 CC disorders, or cancer. The polynucleotide sequences of the invention were
 CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
 CC in some cases, sequences obtained from one or more public databases.
 CC Note: The sequence data for this patent did not form part of the printed

PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231442.
PR 08-SEP-2000; 2000US-0231443.
PR 08-SEP-2000; 2000US-0231444.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
DR WPI; 2001-465570/50.
DR N-PSDB; AAL01353.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen -
XX is used in preventing, treating or ameliorating a medical condition -
XX
XX Claim 11; SEQ ID NO 4041; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a protein of the invention.
XX
SQ Sequence 245 AA;
Query Match 43.8%; Score 716; DB 22; Length 245;
Best Local Similarity 46.2%; Pred. No. 1.5e-51;
Matches 109; Conservative 48; Mismatches 79; Indels 0; Gaps 0;
QY 40 CAECQOLIGHDSRELFVEDRHFHGGCFRCRCORSRLADEPFTQDSLELLCNDYCYSAFSS 99
Db 2 CEECKPIESDSKDLCKDRHWHGECFKCTKNHSLVEKPFPAKDERLLCTECSNECSS 61
QY 100 QSACGETVMPGSRKLEYGGQTWHEHCFLGSGCEQPLGSRSFVDPKGAHYCVPCYENKFA 159
Db 62 KCFCHKRTIMPGSRKMEFGNYWHXTCXVCENCRCQPIXTKPLISKESGNTCVPCXEKEFA 121
QY 160 PRCARCSKTLTGGVTYRQPHRECLVCTGCTQPLAGQOFTSRDEDPYCVACFGELFAP 219
Db 122 HCNFKCKVITSGGITFCQDLWHKEXFLCSGCRDRDCEEFQMFMRDDTFFCXDCYNHLYAN 181
QY 220 KCSSCKRPVGLGGKGYVSPEDRHHNCFSCARCSTSLVGQGFVDPDQVLQCGC 275
Db 182 KCVACSKPISGLTGAKFICFQDSQWHSECFNCCKCSVSLVGKFLQTKNEIFCQKC 237
RESULT 5
AAU20437
ID AAU20437 standard; Protein; 245 AA.
XX
AC AAU20437;
XX
DT 06-DEC-2001 (first entry)
XX

Db 2 CECEKKPIESDSDKDLICYKDRHWHHEGCFKTKCNHSLVLPKFFAAKDBRLCLCTECYSNECSS 61

Qy 100 QCSACGETWPGSRKLEYGGQTWHEHCFLSCGEQPLGSRSFVPDRGAHYCVPCYENKEFA 159
: | : ||||| : | | | | | | | | : : : : ||||| : | |
Db 62 KCFCHKRTIMPGRKMEFKNGYHXTCXVCENCROPITXPKPLISKESGNVCPCXEKEFA 121
: | : ||||| : | | | | | | | | : : : : ||||| : | |
Qy 160 PROCARSKLTQGGVTYRDPQPHRECLVCTGCOTPLAGOOFTSRDEDPYCVCACFGELFAP 219
: | : ||||| : | | | | | | | | : : : : ||||| : | |
Db 122 HYNCFCKVITSGITFCQLWKHKEXFLCSCGCRKDLCEEQFMRSRDDYPFCXDCYNHLAYN 181
: | : ||||| : | | | | | | | | : : : : ||||| : | |
Qy 220 KCSSCKRPVIGLGGGVYVFEDRHWHNCESCARCSTSLVGOGFVDPGDQVLCOGC 275
: | : ||||| : | | | | | | | | : : : : ||||| : | |
Db 182 KVACSKPISGLTGAKFICFDQSQWSECFNCGKSVSLVGKGFLTQNKEIFCQKC 237
: | : ||||| : | | | | | | | | : : : : ||||| : | |

RESULT 6
ABB64785
ID ABB64785 standard; Protein; 245 AA.
XX AC ABB64785;
DT XX
DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 21147.
DE XX
DE Drosophila; developmental biology; cell signalling; insecticide;
KW KW pharmaceutical.
XX OS Drosophila melanogaster.
XX WO200171042-A2.
PN XX
PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL08888.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
PT Disclosure; SEQ ID NO 21147; 2lpp + Sequence Listing; English.
PS CC
CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABE16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 245 AA;

Query Match 38.0%; Score 622; DB 22; Length 245;
Best Local Similarity 50.5%; Pred. No. 9.2e-44;
Matches 100; Conservative 34; Mismatches 52; Indels 12; Gaps 2;
Qy 80 FTQDSSELNCDCYSAFSQCSACGETWPGSRKLEYGGQTWHEHCFLSCGEQPLGSR 139
Db 56 FYCOLGD-----YCRLLGPASKA-----GPKMEYKTRWHENCFCCVCKTAIGTK 103
: | : ||||| : | | | | | | | | : : : : ||||| : | |
: | : ||||| : | | | | | | | | : : : : ||||| : | |

cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Query Match 25.3%; Score 413.5; DB 22; Length 328;
Best Local Similarity 27.5%; Pred. No. 2.3e-26;
Matches 84; Conservative 48; Mismatches 135; Indels 39; Gaps 6;

QY 3 ESFDCAKNESLYGKRYTQDSGYPCYDNTFANTCAEQQLIGHDSRELFYEDRHFH 62
DB 29 EHFLCHHDEQILDATFNQSGPEPCVNCFCFERYTYTCAGCKKPI--LEKTCAMGESWH 86
QY 63 EGCFFRC--RCORSLADEPFTCDQSELNCDYCSAFSSOCAGCTVMPGS----- 112
DB 87 EDCP-CCGACKPLANOTFYERDGKPYCKDYEDLFAARCAKCEKPTDSAVLAMNVKW 145
QY 113 -----RKLEYGGQTWHEHCLVCTG-CQTPLAGQOFTSRDEDPYCV 151
DB 146 HRDCFRNCNIVCHKQEAITRMITALGKTWHEPHEFLCHHDEQILDATFNQSGPEPCV 205
QY 152 PCYENKFAPRCAKSKTLTQGGVYRDQWHEHCLVCTG-CQTPLAGQOFTSRDEDPYCV 210
DB 206 KCFERYTYTCAGCKKPILEKTCAMGERWHEACFCGCGACKKPLASQTFYERDGKPYCK 265
QY 211 ACFGELFAPKSSCKRPIVGLGGKYVSFEDRHHNCFSCARCGSTSLVGGQFVDPGDQV 270
DB 266 QDYENLFAARCAKCEKPI-----DSAVLAMNVKHNCFQCNKENPITSQTFIDGDKP 321
QY 271 LQCGCS 276
DB 322 VCPACN 327

RESULT 9
ABG11960
ID ABG11960 standard; Protein; 116 AA.
XX
AC ABG11960;
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #11951.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS76147.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess biodiversity -
Claim 20; SEQ ID No 42319; 103pp; English.
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.
Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Query Match 21.3%; Score 349; DB 22; Length 116;
Best Local Similarity 48.6%; Pred. No. 1.7e-21;
Matches 52; Conservative 23; Mismatches 32; Indels 0; Gaps 0;

QY 169 LTQGGVYRDQWHEHCLVCTGCTPLAGQOFTSRDEDPYCVACGELFAPKSSCKRPI 228
DB 2 ITSGGITFCDDLWHKEGFLCSGCRKDLCEQFMGRDDYPFCVCYNHLYANKVACSKPI 61
QY 229 VGLGGKYVSPEDRHHNCFSCARCGSTSLVGGQFVDPGDQVLCQGC 275
DB 62 SGLTGAKFICFQDSQWHSECFNCCKSVSLVKGFLTKQNKFCQKC 108

RESULT 10
ABB67068
ID ABB67068 standard; Protein; 746 AA.
XX
AC ABB67068;
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 27996.
XX
KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL11171.
XX


```

XX  Venter JC, Adams M, Li PWD, Myers EW;
XX  WPI; 2001-656860/75.
XX  N-FSDB; ABL11212.
XX
XX  New isolated nucleic acid detection reagent for detecting 1000 or more
XX  genes from Drosophila and for elucidating cell signalling and cell-cell
XX  interactions -
XX
XX  Disclosure; SEQ ID NO 28119; 2lpp + Sequence Listing; English.
XX
XX  The invention relates to an isolated nucleic acid detection reagent
XX  capable of detecting 1000 or more genes from Drosophila. The invention is
XX  useful in developmental biology and in elucidating cell signalling and
XX  cell-cell interactions in higher eukaryotes for the development of
XX  insecticides, therapeutics and pharmaceutical drugs. The invention
XX  discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX  sequences (ABL01840-ABL16175) and the encoded proteins
XX  (ABB57737-ABB72072).
XX  The sequence data for this patent did not form part of the printed
XX  specification, but was obtained in electronic format directly from WIPO
XX  at ftp.wipo.int/pub/published_pct_sequences.
XX
XX  Sequence 497 AA;
XX
XX  Query Match          20.0%; Score 327; DB 22; Length 497;
XX  Best Local Similarity 35.4%; Pred. No. 5,1e-19;
XX  Matches 62; Conservative 27; Mismatches 80; Indels 6; Gaps 2;
XX
XX  QY 101 CSAGETVMPGSRKLEYGGQWHEHCFLCSCGEOPLGSRSPVDPKGAHYCPVCYENKFPAP 160
XX      I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
XX  DB 322 CNACEKPIV--QGVITALGKTWHPHEFTCNHCSQELGTRNFFERDGPYCEPDYHNLFSF 379
XX
XX  QY 161 RCARCSKTLTGGVYTDQPHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGELFAPK 220
XX      I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
XX  DB 380 RCAYCNGAILDKVCTALDKTWTHTHEFFCAQCQQQFGEEGFHERDCKPYCRNDYFEMFAPK 439
XX
XX  QY 221 CSSCKRPVGLGGKYVSEFDRHWHNCFSCARGSTSLVGQGVDPDQVLCQCG 275
XX      I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
XX  DB 440 CNGCNRAIM---ENYISALNSQWHPDCFCVCRDCKAVRGSFYAMEGKPYCPQC 490
XX
XX
XX  RESULT 13
XX  ABB65964
XX  ID ABB65964 standard; Protein; 177 AA.
XX
XX  AC ABB65964;
XX
XX  DT 26-MAR-2002 (first entry)
XX
XX  DE Drosophila melanogaster polypeptide SEQ ID NO 24684.
XX
XX  KW Drosophila; developmental biology; cell signalling; insecticide;
XX  KW pharmaceutical.
XX
XX  OS Drosophila melanogaster.
XX
XX  PN WO200171042-A2.
XX
XX  PD 27-SEP-2001.
XX
XX  XX 23-MAR-2001; 2001WO-US09231.
XX  PF
XX  PR 23-MAR-2000; 2000US-191637P.
XX  PR 11-JUL-2000; 2000US-0614150.
XX
XX  PA (PEKE ) PE CORP NY.
XX
XX  Venter JC, Adams M, Li PWD, Myers EW;
XX  WPI; 2001-656860/75.
XX  N-FSDB; ABL10067.

```

```

XX  New isolated nucleic acid detection reagent for detecting 1000 or more
XX  genes from Drosophila and for elucidating cell signalling and cell-cell
XX  interactions -
XX
XX  Disclosure; SEQ ID NO 24684; 2lpp + Sequence Listing; English.
XX
XX  The invention relates to an isolated nucleic acid detection reagent
XX  capable of detecting 1000 or more genes from Drosophila. The invention is
XX  useful in developmental biology and in elucidating cell signalling and
XX  cell-cell interactions in higher eukaryotes for the development of
XX  insecticides, therapeutics and pharmaceutical drugs. The invention
XX  discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX  sequences (ABL01840-ABL16175) and the encoded proteins
XX  (ABB57737-ABB72072).
XX  The sequence data for this patent did not form part of the printed
XX  specification, but was obtained in electronic format directly from WIPO
XX  at ftp.wipo.int/pub/published_pct_sequences.
XX
XX  Sequence 177 AA;
XX
XX  Query Match          19.9%; Score 325.5; DB 22; Length 177;
XX  Best Local Similarity 45.0%; Pred. No. 2.3e-19;
XX  Matches 54; Conservative 25; Mismatches 38; Indels 3; Gaps 1;
XX
XX  QY 2 SESEFDCARCNESLYGRKYIQTDSGPYCPVCYDNTFANTCACCQQLIGHDSRELFVEDRRF 61
XX      I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
XX  DB 60 SGHCCWCQDESILTGQRVIRDDHPYCIKCYENFANTCEENKIIGIDSKDLVKDKHW 119
XX
XX  QY 62 HEGFRRCRCORSRLADEPFTQDSLELNCDCYSAFSSQCSACGETVMPGSRKLEYGGQT 121
XX      I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
XX  DB 120 HEACFLCFKCHLSLVDFQFGAKADKIYGCNGCYDAQFASRCDCGCGEVFRAGKQIL---GET 176
XX
XX
XX  RESULT 14
XX  AAW88807
XX  ID AAW88807 standard; Protein; 292 AA.
XX
XX  AC AAW88807;
XX
XX  DT 01-MAR-1999 (first entry)
XX
XX  DE Polypeptide fragment encoded by gene 56.
XX
XX  KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX  KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX  KW developmental abnormality; foetal deficiency; blood; allergy; renal;
XX  KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX  KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX  KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX  KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX  KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX  OS Homo sapiens.
XX
XX  PN WO9854963-A2.
XX
XX  PD 10-DEC-1998.
XX
XX  XX 04-JUN-1998; 98WO-US11422.
XX  PF
XX  PR 18-DEC-1997; 97US-0070923.
XX  PR 06-JUN-1997; 97US-0048877.
XX  PR 06-JUN-1997; 97US-0048881.
XX  PR 06-JUN-1997; 97US-0048884.
XX  PR 06-JUN-1997; 97US-0048893.
XX  PR 06-JUN-1997; 97US-0048896.
XX  PR 06-JUN-1997; 97US-0048899.
XX  PR 06-JUN-1997; 97US-0048915.
XX  PR 06-JUN-1997; 97US-0048949.
XX  PR 06-JUN-1997; 97US-0048964.
XX  PR 06-JUN-1997; 97US-0048972.
XX  PR 06-JUN-1997; 97US-0049020.

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:56:17 ; Search time 29 seconds
(without alignments)
408.519 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSESFDCARCNESLYGRKYI.....QGFVDPGDQVLCQCSQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	314	19.2	292	4	US-09-205-258-667
2	314	19.2	334	4	US-09-205-258-672
3	306.5	18.7	314	1	US-08-269-441A-2
4	306.5	18.7	314	3	US-09-008-465-3
5	306.5	18.7	314	3	US-09-276-851-2
6	306.5	18.7	314	4	US-09-528-959-3
7	304.5	18.6	591	2	US-08-889-402-1
8	304.5	18.6	605	2	US-08-889-402-2
9	297.5	18.2	341	3	US-09-008-465-1
10	297.5	18.2	341	4	US-09-528-959-1
11	293	17.9	454	1	US-08-166-316-2
12	293	17.9	457	4	US-09-124-238A-10
13	293	17.9	457	4	US-09-721-975-10
14	293	17.9	457	4	US-09-986-621-10
15	284	17.4	457	4	US-09-124-238A-1
16	284	17.4	457	4	US-09-721-975-1
17	284	17.4	457	4	US-09-986-621-1
18	262	16.0	612	4	US-08-894-454-163
19	257	15.7	343	4	US-08-894-454-161
20	220	13.5	421	3	US-09-002-567B-1
21	220	13.5	421	3	US-09-002-567B-3
22	220	13.5	421	4	US-09-571-347-1
23	220	13.5	421	4	US-09-571-347-3
24	200	12.2	194	2	US-08-739-485-8
25	197.5	12.1	194	1	US-08-616-368A-11
26	197.5	12.1	194	3	US-09-054-298-11
27	197.5	12.1	194	3	US-08-818-655-11

Sequence 11, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 8, Appl
Sequence 13, Appl
Sequence 9, Appl
Sequence 8, Appl
Sequence 13, Appl
Sequence 8, Appl
Sequence 13, Appl
Sequence 8, Appl
Sequence 13, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl

28 197.5 12.1 194 4 US-09-305-839-11
29 196.5 12.0 193 1 US-08-616-368A-1
30 196.5 12.0 193 2 US-08-739-485-1
31 196.5 12.0 193 3 US-09-054-298-1
32 196.5 12.0 193 3 US-08-818-655-1
33 196.5 12.0 193 4 US-09-305-839-1
34 195.5 12.0 193 1 US-08-616-368A-8
35 195.5 12.0 193 1 US-08-616-368A-13
36 195.5 12.0 193 2 US-08-739-485-9
37 195.5 12.0 193 3 US-09-054-298-8
38 195.5 12.0 193 3 US-09-054-298-13
39 195.5 12.0 193 3 US-08-818-655-8
40 195.5 12.0 193 3 US-08-818-655-13
41 195.5 12.0 193 4 US-09-305-839-8
42 195.5 12.0 193 4 US-09-305-839-13
43 182 11.1 193 1 US-08-616-368A-9
44 182 11.1 193 3 US-09-054-298-9
45 182 11.1 193 3 US-08-818-655-9

ALIGNMENTS

RESULT 1

US-09-205-258-667
; Sequence 667, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06


```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0460 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 314 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1516012
;
US-09-008-465-3
Query Match 18.7%; Score 306.5; DB 3; Length 314;
Best Local Similarity 27.4%; Pred. No. 2.8e-21;
Matches 77; Conservative 40; Mismatches 145; Indels 19; Gaps 7;

QY 3 ESFDCAKNESLYGRKYIQTDSGPYCPYNTFANTCAEQQLIGHDSRELFDH 62
D 34 QCFVCAQCFOQFPPEGLFVEFEGRKYCEHDFQMLFAPCCHQCGETI--IGRVIKAMNNSWH 91
QY 63 EGCPRCCRCQSLADEPTCQDSSELLCNDYCYSAFSSQ-----CSACGETVMPGSRKLEY 117
D 92 PECFRCDLCQEVLDIGFVKNAGRHLRCRPNHREKARGLGKGYCOCKHAI--DEQPLIF 149
QY 118 GGQTWHEHCFCLSCGEQPL--GSRFVDPDKGAHYCVPCYENKFAPRCARCSKTLTQGGVT 175
D 150 KNDPYHPDHFNCANGKELTADAREL---KGELYCLPCHDKMGVPCICACRRPIEGRVN 206
QY 176 YRDQPHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGEFAPKCSSCKRPVIGLGGK 235
D 207 AMGQWVHEHFVCAKCEKPFGLGHRHYERKGLAYCETHYNQLFGDVCFFHCNRYI----EGD 262
QY 236 YVSFEDRHHNCFSCARCSL--VGQGFVDPDQGVLCQGC 275
D 263 VVSALNKAWCNCVFACSTCNTKLTILKNKFEVDFDKMPVCKKC 303

```

```

RESULT 5
US-09-276-851-2
; Sequence 2, Application US/09276851
; Patent No. 6245522
; GENERAL INFORMATION:
; APPLICANT: Rearden, Ann
; TITLE OF INVENTION: A MASTER MOLECULAR RHEOSTAT SWITCH FOR
; TITLE OF INVENTION: CELL SIGNALING
; FILE REFERENCE: REGEN1440
; CURRENT APPLICATION NUMBER: US/09/276,851
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-276-851-2
Query Match 18.7%; Score 306.5; DB 3; Length 314;
Best Local Similarity 27.4%; Pred. No. 2.8e-21;
Matches 77; Conservative 40; Mismatches 145; Indels 19; Gaps 7;

QY 3 ESFDCAKNESLYGRKYIQTDSGPYCPYNTFANTCAEQQLIGHDSRELFDH 62
D 34 QCFVCAQCFOQFPPEGLFVEFEGRKYCEHDFQMLFAPCCHQCGETI--IGRVIKAMNNSWH 91
QY 63 EGCPRCCRCQSLADEPTCQDSSELLCNDYCYSAFSSQ-----CSACGETVMPGSRKLEY 117

```

```

D 92 PECFRCDLCQEVLDIGFVKNAGRHLRCRPNHREKARGLGKGYCOCKHAI--DEQPLIF 149
QY 118 GGQTWHEHCFCLSCGEQPL--GSRFVDPDKGAHYCVPCYENKFAPRCARCSKTLTQGGVT 175
D 150 KNDPYHPDHFNCANGKELTADAREL---KGELYCLPCHDKMGVPCICACRRPIEGRVN 206
QY 176 YRDQPHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGEFAPKCSSCKRPVIGLGGK 235
D 207 AMGQWVHEHFVCAKCEKPFGLGHRHYERKGLAYCETHYNQLFGDVCFFHCNRYI----EGD 262
QY 236 YVSFEDRHHNCFSCARCSL--VGQGFVDPDQGVLCQGC 275
D 263 VVSALNKAWCNCVFACSTCNTKLTILKNKFEVDFDKMPVCKKC 303

```

```

RESULT 6
US-09-528-959-3
; Sequence 3, Application US/09528959
; Patent No. 6379904
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Corley, Neil C.
; TITLE OF INVENTION: HUMAN PINCH PROTEIN HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
;
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/528,959
FILING DATE: 20-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/008,465
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0460 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 314 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1516012
;
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-528-959-3
Query Match 18.7%; Score 306.5; DB 4; Length 314;
Best Local Similarity 27.4%; Pred. No. 2.8e-21;
Matches 77; Conservative 40; Mismatches 145; Indels 19; Gaps 7;

QY 3 ESFDCAKNESLYGRKYIQTDSGPYCPYNTFANTCAEQQLIGHDSRELFDH 62
D 34 QCFVCAQCFOQFPPEGLFVEFEGRKYCEHDFQMLFAPCCHQCGETI--IGRVIKAMNNSWH 91
QY 63 EGCPRCCRCQSLADEPTCQDSSELLCNDYCYSAFSSQ-----CSACGETVMPGSRKLEY 117

```

Db 92 PECFRCDLQCEVLADIGFVKNAGRLRCPCHNREKARGLGKVICQKCHAI--DEOPLIF 149
QY 118 GGQTHWEHCFCLSCGQPL--GSRFVPDKGAHYCPVCYENKFAPCARCCKTTLTGGVY 175
Db 150 KNDPYHDFHNCANCKELTADREL---KGELYCLPCHDKMGVPCGACRRPIEGRVYN 206
QY 176 YRDQPHRECLVCTGQOTPLAQOFTSRDEDPYCVACFGELFAPKCSSCKRPVIGLGQK 235
Db 207 AMGKOWHVEHFVCAKCEKFLGHRHYERKGLAYCETHYNQLFGDVCFHCRNVI----EGD 262
QY 236 YVSFDRHHNCFSCARCSTL-VGQGFVPDGDQVLCQC 275
Db 263 VVSALNKACVNCVFACSTCNTLTLLANKFEVDMKPVCKK 303

RESULT 7

US-08-889-402-1
; Sequence 1, Application US/08889402
; Patent No. 5811288
; GENERAL INFORMATION:
; TITLE OF INVENTION: NOVEL ISOFORM GENE FOR FOCAL ADHESION
; NUMBER OF SEQUENCES: 7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/889,402
; FILING DATE:

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 178334/1996
FILING DATE: 08-JUL-1996
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Monoblastic cell and placenta
CELL LINE: U937
US-08-889-402-1

Query Match 18.6%; Score 304.5; DB 2; Length 591;
Best Local Similarity 27.5%; Pred. No. 9e-21;
Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3;

QY 40 CAECQQLIGHDSRELFYEDRHFEHGCRCRCQRSRLADEPFTQDSELLCNDYCYSAFSS 99
Db 358 CGACKKPAA--GOVVTAMGKTWHPHFVCTHCOEEIGSRNFFRDQGPCEKDYHNLFS 415
QY 100 QCSACGETVMPGSRKLEYGGQTHWEHCFCLSCGCEQPLGSRSFVPDKGAHYCPVCYENKFA 159
Db 416 RCYICNGPIL--DKVVTALDRTWHPHFVCAQCGAFFGPGHEKDGKAYCRKDYDMEA 473
QY 160 PRCARCKTLTGGVTVYRDPQPHRECLVCTGQOTPLAQOFTSRDEDPYCVACFGELFAP 219
Db 474 PKGGCARAILENYISALNTLWHPFCVCRECFPFVNGSFFHHDGQPCVEVHYHRRGS 533
QY 220 KCSSCKRPVIGGGKYVSFEDRHHNCFSCARCSTSLVGQGFVPDGDQVLCQC 275
Db 534 LCSGCKQKPIIT---GRCITAMAKKFPHEFVCAFLKQLNKGTFKQNDKPYCQNC 585

RESULT 8

US-08-889-402-2
; Sequence 2, Application US/08889402
; Patent No. 5811288
; GENERAL INFORMATION:

; TITLE OF INVENTION: NOVEL ISOFORM GENE FOR FOCAL ADHESION
; NUMBER OF SEQUENCES: 7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/889,402
; FILING DATE:

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 178334/1996
FILING DATE: 08-JUL-1996
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Monoblastic cell and placenta
CELL LINE: U937
US-08-889-402-2

Query Match 18.6%; Score 304.5; DB 2; Length 605;
Best Local Similarity 27.5%; Pred. No. 9.2e-21;
Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3;
QY 40 CAECQQLIGHDSRELFYEDRHFEHGCRCRCQRSRLADEPFTQDSELLCNDYCYSAFSS 99
Db 372 CGACKKPAA--GOVVTAMGKTWHPHFVCTHCOEEIGSRNFFRDQGPCEKDYHNLFS 429
QY 100 QCSACGETVMPGSRKLEYGGQTHWEHCFCLSCGCEQPLGSRSFVPDKGAHYCPVCYENKFA 159
Db 430 RCYICNGPIL--DKVVTALDRTWHPHFVCAQCGAFFGPGHEKDGKAYCRKDYDMEA 487
QY 160 PRCARCKTLTGGVTVYRDPQPHRECLVCTGQOTPLAQOFTSRDEDPYCVACFGELFAP 219
Db 488 PKGGCARAILENYISALNTLWHPFCVCRECFPFVNGSFFHHDGQPCVEVHYHRRGS 547
QY 220 KCSSCKRPVIGGGKYVSFEDRHHNCFSCARCSTSLVGQGFVPDGDQVLCQC 275
Db 548 LCSGCKQKPIIT---GRCITAMAKKFPHEFVCAFLKQLNKGTFKQNDKPYCQNC 599

RESULT 9

US-09-008-465-1
; Sequence 1, Application US/09008465
; Patent No. 6174702
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN PINCH PROTEIN HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/008,465
; FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/008,465
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0460 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 341 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SEMVNOT04
CLONE: 3540806
US-09-008-465-1

Query Match 18.28; Score 297.5; DB 3; Length 341;
Best Local Similarity 27.28; Pred. No. 2.2e-20;
Matches 79; Conservative 42; Mismatches 138; Indels 31; Gaps 9;
Qy 5 FDCAK-----NESLY-----GRKYIOTDSGPVPCYDNTFANTCAEQOOLIGHDSRELFY 56
Db 41 FVCAQCFRPPEGLFYEFEGKCYEHD-----FQMLFAPCCGSGEFT--IGRVKA 90
Qy 57 EDRHFEGCFRCRCQSLADEPTCQDSSELLCNDYCASFSSQ-----CSACGETVMPG 111
Db 91 MNNWHPGCFRCELCDELADLGFVNAGRHLCRPNRREKAKGLGKYICQCHLVI--D 148
Qy 112 SRLEYGGQIWHCHFLCSCGEQPLGSRSFVPDKGAHYCVPCYENKFPAPRCARSKLTQ 171
Db 149 EQPLMFSDAYPHDFNCHCGKELTAEA--RELKGLYCLPCHDKMGVPCGACRRPIEG 207
Qy 172 GGVTYRQDPWHRREGLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKSCSKRPVGL 231
Db 208 RVNVALGQWHEHFVCAKCEKPFGLGHRHYEKKGLAYCETHYNQLFGDVCYNCSHVI--- 264
Qy 232 GGGYVVFEDRHHWHNCFSCARCSTSL-VGGGFVPDGDQVLCQCSQAGP 280
Db 265 -EGDVSALNAKAWCVSCFSCSTCNSKLTLNKMFVFDMPKPKCKRYEKFP 313

RESULT 10
US-09-528-959-1
Sequence 1, Application US/09528959
Patent No. 6379904
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
Guegler, Karl J.
Corley, Neil C.
TITLE OF INVENTION: HUMAN PINCH PROTEIN HOMOLOG
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/528,959
FILING DATE: 20-Mar-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/008,465
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0460 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 341 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SEMVNOT04
CLONE: 3540806
US-09-528-959-1

Query Match 18.28; Score 297.5; DB 4; Length 341;
Best Local Similarity 27.28; Pred. No. 2.2e-20;
Matches 79; Conservative 42; Mismatches 138; Indels 31; Gaps 9;
Qy 5 FDCAK-----NESLY-----GRKYIOTDSGPVPCYDNTFANTCAEQOOLIGHDSRELFY 56
Db 41 FVCAQCFRPPEGLFYEFEGKCYEHD-----FQMLFAPCCGSGEFT--IGRVKA 90
Qy 57 EDRHFEGCFRCRCQSLADEPTCQDSSELLCNDYCASFSSQ-----CSACGETVMPG 111
Db 91 MNNWHPGCFRCELCDELADLGFVNAGRHLCRPNRREKAKGLGKYICQCHLVI--D 148
Qy 112 SRLEYGGQIWHCHFLCSCGEQPLGSRSFVPDKGAHYCVPCYENKFPAPRCARSKLTQ 171
Db 149 EQPLMFSDAYPHDFNCHCGKELTAEA--RELKGLYCLPCHDKMGVPCGACRRPIEG 207
Qy 172 GGVTYRQDPWHRREGLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKSCSKRPVGL 231
Db 208 RVNVALGQWHEHFVCAKCEKPFGLGHRHYEKKGLAYCETHYNQLFGDVCYNCSHVI--- 264
Qy 232 GGGYVVFEDRHHWHNCFSCARCSTSL-VGGGFVPDGDQVLCQCSQAGP 280
Db 265 -EGDVSALNAKAWCVSCFSCSTCNSKLTLNKMFVFDMPKPKCKRYEKFP 313

RESULT 11
US-08-166-316-2
Sequence 2, Application US/08166316
Patent No. 5504192
GENERAL INFORMATION:
APPLICANT: Gill, Gordon N.
APPLICANT: Wu, Rui-Yun
TITLE OF INVENTION: ENDOCYTIC CODE BINDING PROTEIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/166,316
FILING DATE: 13-DEC-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 20:00:05 ; Search time 57 Seconds
(without alignments)
648.233 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSEFDCAKNESLYGRKYI.....QGFPDGDQVLCGCSQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	716	43.8	245	11	US-09-764-891-4041
2	314	19.2	232	15	US-10-023-282-667
3	314	19.2	334	15	US-10-023-282-672
4	313	19.1	386	10	US-09-211-424-2
5	306.5	18.7	314	9	US-09-880-705-2
6	306.5	18.7	314	15	US-10-092-0668-3
7	301	18.4	266	15	US-10-043-487-245
8	301	18.4	596	15	US-10-205-823-222
9	297.5	18.2	341	15	US-10-092-0668-1
10	293	17.9	457	11	US-09-986-625-10
11	293	17.9	498	9	US-09-925-302-475
12	284	17.4	457	11	US-09-986-625-1
13	283	17.3	835	15	US-10-273-680-7
14	275	16.8	831	15	US-10-273-680-2
15	251.5	15.4	300	15	US-10-102-806-684

16	232.5	14.2	547	10	US-09-779-307-12	Sequence 12, Appl
17	232.5	14.2	547	10	US-09-779-307-13	Sequence 13, Appl
18	232	14.2	538	10	US-09-779-307-2	Sequence 2, Appl
19	232	14.2	538	10	US-09-779-307-11	Sequence 11, Appl
20	220	13.5	421	12	US-10-357-627-1	Sequence 1, Appl
21	220	13.5	421	12	US-10-357-627-3	Sequence 3, Appl
22	214.5	13.1	663	15	US-10-147-026-4	Sequence 4, Appl
23	214.5	13.1	663	15	US-09-773-926-17	Sequence 17, Appl
24	200	12.2	194	9	US-10-126-099-8	Sequence 8, Appl
25	196.5	12.0	193	14	US-10-126-099-1	Sequence 1, Appl
26	195.5	12.0	193	14	US-10-126-099-9	Sequence 9, Appl
27	190	11.6	159	9	US-09-814-122-54	Sequence 54, Appl
28	179.5	11.0	3075	10	US-09-938-275-5	Sequence 5, Appl
29	178	10.9	193	14	US-10-126-099-7	Sequence 7, Appl
30	164	10.0	401	11	US-09-932-367A-14	Sequence 14, Appl
31	163.5	10.0	398	11	US-09-932-367A-6	Sequence 6, Appl
32	163	10.0	367	11	US-09-932-367A-20	Sequence 20, Appl
33	162	9.9	615	10	US-09-862-027-49	Sequence 49, Appl
34	161	9.8	638	10	US-09-862-027-41	Sequence 36148, A
35	160.5	9.8	536	9	US-09-864-761-36148	Sequence 143, App
36	158.5	9.7	305	10	US-09-771-161A-143	Sequence 142, App
37	158.5	9.7	409	10	US-09-771-161A-142	Sequence 233, App
38	158.5	9.7	647	10	US-09-771-161A-234	Sequence 234, App
39	158.5	9.7	647	10	US-09-932-367A-5	Sequence 5, Appl
40	157.5	9.6	394	11	US-09-932-367A-2	Sequence 2, Appl
41	156.5	9.6	383	11	US-09-932-367A-4	Sequence 4, Appl
42	156.5	9.6	395	11	US-09-932-367A-10	Sequence 10, Appl
43	155.5	9.5	397	11	US-09-932-367A-12	Sequence 12, Appl
44	155.5	9.5	402	11	US-09-932-367A-12	Sequence 2, Appl
45	155.5	9.5	1548	15	US-10-180-903-2	

ALIGNMENTS

RESULT 1

US-09-764-891-4041
; Sequence 4041, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764, 891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 4041
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (86)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (99)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (116)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (147)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (172)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-891-4041

RESULT 3
US-10-023-282-672
; Sequence 672, Application US/10023282
; Publication No. US2003005283A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,891
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 672
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-023-282-672

Query Match 19.2%; Score 314; DB 15; Length 334;
Best Local Similarity 32.8%; Pred. No. 1.4e-20;
Matches 62; Conservative 26; Mismatches 83; Indels 18; Gaps 4;

QY 101 CSACGETVMPGSRKLEYGGTWHCHFLCGCEQPIGSGFVDPDKAHYCVPCYENKFP 160
Db 100 CASCQKPI--AGKVIHALGQSWHPEHFVCTHCKEEIGSSPFFERSGLXXCPNDYHOLFSP 157
QY 161 RCARCSKTLTQGGVYTDQPHWRECLVCTGCTPLAGQQFTSRDEDPYCVACFGELFAPK 220
Db 158 RCAYCAAPILDKVLTMNQTWHPHFHFCSHCGEVFGAEGFHEKDKPKYCKRDKFLAMFSPK 217
QY 221 CSSCKRPYVGLGGKYVSFEDRHHHNCFCARCSTSL-VGGGFVDPGDO----- 269
Db 218 CGGCNRPVL---ENYLSAMDTVMHPECFVCGDCFTSFSTGTSFFELDGRPFCELHYHRR 273
QY 270 -VLCQGCSCQ 277
Db 274 GTLCHGCGQ 282

RESULT 4
US-09-211-424-2
; Sequence 2, Application US/09211424A
; Patent No. US20020177231A1
; GENERAL INFORMATION:
; APPLICANT: Staunton et al.
; TITLE OF INVENTION: Leupaxin Materials and Methods
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/211,424A
; CURRENT FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-211-424-2

Query Match	19.18;	Score 313;	DB 10;	Length 386;
Best Local Similarity	32.88;	Pred. No. 2.1e-20;		
Matches	62;	Conservative 26;	Mismatches 83;	Indels 18; Gaps 4
Qy	101	CSACGEVTMPGSRRLKEYGGGTWHEHCFCLSCGEOPGLGSRSVFPDKGAHYCVPCYENKFAF	160	
Db	152	CASCQKPI--AGKVIHALGQSWHPEHFVCTHCKEEIGSSPFFERSGLAYCNDYHQLESP	209	
Qy	161	RCARCSYTLTQGGVYTRDQPHWRRELVCTGCGQTPLAGQQFTSRDEDDPYCVACGELFAPK	220	
Db	210	RCAYCAAPILDKVLVTANQTHWPEHFFCSHCGEVFGAEGFHEKDKKPYCRKDFLAMEFSPK	269	
Qy	221	CSSCKRPVLGLGGKKYVSFEDRWHHNCFCARCSTSL-VSQGFVPDGDQ-----	269	
Db	270	CGGCGNRPLV----ENILSAMDTVWHPECFVCGDCFTSFSTGSPFELDGRPRFCELYHHR	325	
Qy	270	-VLQCGCSQ	277	
Db	326	GTLCGCGQ	334	

```

RESULT 5
US-09-880-705-2
; Sequence 2, Application US/09880705
; Patent No. US20020081639A1
; GENERAL INFORMATION:
; APPLICANT: Rearden, Ann
; TITLE OF INVENTION: A MASTER MOLECULAR RHEOSTAT SWITCH FOR
; FILE OF INVENTION: CELL SIGNALING
; FILE REFERENCE: REGEN1440
; CURRENT APPLICATION NUMBER: US/09/880,705
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 09/276,851
; PRIOR FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-705-2

```

[illegible]

RESULT 6
US-10-092-066B-3
; Sequence 3, Application US/10092066B
; Publication No. US20030104472A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.

Db 90 CAHCNQVIR--GPFVLALGKSWHPPEFNCAHCNKTWAYIGFVEKGALYCELCEYKFFAP 147
QY 161 RCARCSKTLTQGGVYTRDQPHRECLVCTGCTQPLAGCQOFTSRDEDPYCVACFGELFAPK 220
Db 148 ECGRCQKILGEVINALKQTHVHSCFVACGKPIRNNVPHLEDEGEPCYCTDYALFGTI 207
QY 221 CSSCKRPVIGLGGKYYFFEDRHHNCFSCARCSTSLVQGGFVDPDQVLCQ 273
Db 208 CHGCEPPIE--AGDMFLEALGYTHDTCFCVSCVCCESLEGQTFFSKKDKPLCK 258

RESULT 8

US-10-205-823-222
; Sequence 222, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-222

Query Match 18.4%; Score 301; DB 15; Length 596;
Best Local Similarity 34.1%; Pred. No. 3.9e-19;
Matches 59; Conservative 21; Mismatches 89; Indels 4; Gaps 2;
QY 101 CSACGETVMPGSRKLEYGGTWHCHFLCSCGCEQPLGSRFVDPDKGAHYVPCYENKFP 160
Db 420 CAHCNQVIR--GPFVLALGKSWHPPEFNCAHCNKTWAYIGFVEKGALYCELCEYKFFAP 477
QY 161 RCARCSKTLTQGGVYTRDQPHRECLVCTGCTQPLAGCQOFTSRDEDPYCVACFGELFAPK 220
Db 478 ECGRCQKILGEVINALKQTHVHSCFVACGKPIRNNVPHLEDEGEPCYCTDYALFGTI 537
QY 221 CSSCKRPVIGLGGKYYFFEDRHHNCFSCARCSTSLVQGGFVDPDQVLCQ 273
Db 538 CHGCEPPIE--AGDMFLEALGYTHDTCFCVSCVCCESLEGQTFFSKKDKPLCK 588

RESULT 9

US-10-092-066B-1
; Sequence 1, Application US/10092066B
; Publication No. US20030104472A1

; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: ANTIBODY SPECIFICALLY BINDING HUMAN PINCH PROTEIN HOMOLOG
; FILE REFERENCE: PF-0460-2CIP
; CURRENT APPLICATION NUMBER: US/10/092,066B
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/528,959
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 09/008,465
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030104472A1 3540806CD1
US-10-092-066B-1
Query Match 18.2%; Score 297.5; DB 15; Length 341;
Best Local Similarity 27.2%; Pred. No. 4.5e-19;
Matches 79; Conservative 42; Mismatches 138; Indels 31; Gaps 9;
QY 5 FDCAKC---NESLY-----GRKYIOTDSGGYVPCYDNTTANTCAEQQLIGHDSRELFY 56
Db 41 FVCAQCPRPPEGLFYEFEGKRYCEHD-----FQMLFAPCCGCGEFT--IGRVKA 90
QY 57 EDRHFHEGCFRCRCQRSLADEPFTQDSELLCNDYCYSASFSSQ-----CSACGETVMPG 111
Db 91 MNNWHPGCFRCELCDVELADLGFVNAGRHLRCPCHNREKAKGLGYICORCHLVI--D 148
QY 112 SRKLEYGGTWHCHFLCSCGCEQPLGSRFVDPDKGAHYVPCYENKFPACRASKTLTQ 171
Db 149 EQLMFRSDAYHPDFNCTHCGKELTAEA-RELKGLYCLPCHDKMGVPICGACRRPIEG 207
QY 172 GGVYTRDQPHRECLVCTGCTQPLAGCQOFTSRDEDPYCVACFGELFAPKSSCKRPVGL 231
Db 208 RVVNALGQWHEHFVCAKCEKPLGHRHHEKGLAYCETHYNQLFGDVYCNCSHVI--- 264
QY 232 GGGKYVSFEDRHHNCFSCARCSTSL-VGQGFVDPDQVLCQCSQAGP 280
Db 265 -EGDVVSALNKAWCVSCFCSTCNSKLTLNKFEVFDMPKVCRCYEKFP 313

RESULT 10

US-09-986-625-10
; Sequence 10, Application US/09986625
; Publication No. US20030125248A1
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. US20030125248A1el Bone Mineralization Proteins, DNA, Vecto
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/986,625
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 09/124,238
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 10
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-986-625-10

	Query Match	17.9%	Score 293;	DB 11;	Length 457;
	Best Local Similarity	32.4%;	Pred. No. 1.6e-18;		
	Matches	56;	Conservative 27;	Mismatches 86;	Indels 4; Gaps 2;
QY	101 CSAGETVMPGSRKLEVGCGTWHCHFLCSGCCEQPLGSRFVPDKGAHYCVCPCYENFAP	160			
Db	: : : : : : : : : :				
282 CHQCHKVIR--GRYLVALGHAYHPPEEFCVQCCKGLVEGGFFEEKGAIFCCPCYDVRYP	339				
QY	161 RCARCSKTLTGQGVTVYRDQPWHRDECLVCTGCTPLAQOQTSTRDEDVVCACFCELPAKP	220			
Db	: : : : : : : : :				
340 SCACKKKITGEIMHALKWTHHHVHCFTCAACKTPINRAFYMEGVYPCERDYEKMFGTK	399				
QY	221 CSSCKRPIVLGGGGKYVSFEDRRHWHNCFSCARCSTSLVGQGFVPDGDQLCQ	273			
Db	400 CHGCDPFKID--AGDRFTEALGFSWHDPCFVCAICQINLEGRTFFYSKKDRPLCK	450			

```

; Publication No. US20030125248A1
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: NO. US20030125248A1el Bone Mineralization Proteins, DNA, Vecto
; TITLE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/986,625
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 09/124,238
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 1
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-986-625-1

Query Match          17.4%; Score 284; DB 11; Length 457;
Best Local Similarity 31.8%; Pred. No. 1e-17;
Matches 55; Conservative 27; Mismatches 87; Indels 4; Gaps 2;

Qy   101 CSACGETVMPGSKRLYGQGTTWHEHCFILSGCEQPLGSRSFVPDKGAHYCVPCYENKTPAP 160
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db   282 CHQCHKIIR--GRYLVALGHAYHPBEFVSQCCKVLEGGFFFEERGAIFCPSCYDVRYAP 339
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Qy   161 RCARCKSTLTQGGVTYRDQPWHRECLVCTGCOTPLAGQQFTSRDEDPYVCACFGBLEFAPK 220
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db   340 SCACKCKKTGTETIMHALKMTHVWPCTCAACKTPTINRAFYWEEGCAPYCERYEKMGTK 399
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Qy   221 CSSCKRPIVLGCGKKYVSFEDRRHHNHCSCARCSTSLVGQGFVPGDQVLQ 273
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db   400 CRCDFDKID--ADGREALEALGFSWHDTCFWCAICQINLEGKFYFKKKDKPLCK 450
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |

```

RESULT 12
US-09-986-625-1
; Sequence 1, A

```
Db 98 -----QMSAQRKKEALGRGNIKMLSRVHMATCEKCEKINGGEVAIFVSRAGPGVC 150
QY 122 WHEHCFCLSCGEQPLGSRSPVDPKGAHYCPVCYENKFAPCARCSTLTQGGVYRD-QP 180
Db 151 WHSCFVCSTCNELLDVLIYFYQDGKIHGRHHAELLKPCRSACDEIIFADECTEAEGRH 210
QY 181 WHEHRECLVCTGQTPLAGQFTSRDEDPYCVACFGELFAPKCSSCKRPIVGLGGKYSFE 240
Db 211 WHMHFCCYBCTETVLGGQRYIMKDRPFCCGCGSHYAEYCSCGEHI-GVDHAQ-MTYD 268
QY 241 DRWHH--HNCFCARCSTSLVGQGFVPDGDQV--LCOGCS 276
Db 269 GQHWATETCFSCAQCKVSLGCPFLPKRGRIYCKRACS 307
```

RESULT 14

```
US-10-273-680-2
; Sequence 2, Application US/10273680
; Publication No. US2003009996A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Tomlinson, James E.
; APPLICANT: Komuves, Laszlo G.
; TITLE OF INVENTION: VEP2, A VASCULAR ENDOTHELIAL CELL
; TITLE OF INVENTION: SPECIFIC AND LIM DOMAIN CONTAINING MOLECULE AND USES
; FILE REFERENCE: THEREFOR
; FILE REFERENCE: MPI01-317PIRNM
; CURRENT APPLICATION NUMBER: US/10/273,680
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: US 60/329,756
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-273-680-2
```

```
Query Match 16.8%; Score 275; DB 15; Length 831;
Best Local Similarity 27.0%; Pred. No. 1.2e-16;
Matches 76; Conservative 43; Mismatches 106; Indels 56; Gaps 13;

QY 28 CVPYDNTFANTCAEQQLIGHDSRELFYEDRHFEGCFRCRCORSLADEPFTCODSEL 87
Db 51 CLPEERKVPVNSPGE-----KHRIKQLLYQ-LPPHNEVRYC---QSLSEE-----EKKEL 97
QY 88 LCMDCYCSAFSSQ-----CSAC-----GETVMPGSRKLEYGG 119
Db 98 -----QVFSQRKKEALGRGTIKLLSRVHMHAVCEQCGLKINGGEVAVFASR--AGPG 148
QY 120 QTWHEHCFCLSCGEQPLGSRSPVDPKGAHYCPVCYENKFAPCARCSTLTQGGVYRD- 178
Db 149 VCMHPSCFVCFCTCNELLDVLIYFYQDGKIHGRHHAELLKPCRSACDEIIFADECTEAE 208
QY 179 QPWHRECLVCTGQTPLAGQFTSRDEDPYCVACFGELFAPKCSSCKRPIVGLGGKYSVS 238
Db 209 RWHMHKHFCCLECETVLGGQRYIMKDRPFCCGCGSHYAEYCETGEHI-GVDHAQ-MT 266
QY 239 FEDRHH--HNCFCARCSTSLVGQGFVPDGDQVLC--QGCS 276
Db 267 YDQHWATEACFSCAQCKASLIGCPFLPKQGIYCKTCS 307
```

RESULT 15

```
US-10-102-806-684
; Sequence 684, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103PIC1
```

```
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 684
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-806-684

Query Match 15.4%; Score 251.5; DB 15; Length 300;
Best Local Similarity 28.0%; Pred. No. 5.5e-15;
Matches 61; Conservative 25; Mismatches 97; Indels 35; Gaps 5;

QY 69 CRCORSIADBPFTTCQDSSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQTHHEHCF 128
Db 106 CGCQTLRPM-----STPGGAAAVRALGQLFHIACFT 139
QY 129 CSCEQPLGSRSPVDPKGAHYCPVCYENKFAPCARCSTLTQGGVYRDQPHRECLVC 188
Db 140 CHCAQOLQGOQFYSLGAPYCEGYDTDL-EKNTCGEPIITDMLRATGKAYHPHCFTC 198
QY 189 TGCQTPLAGQOF-TSRDEDPYCVACFGELFAPKCSSCKRPIV---GLGGKYSVFEDRHW 244
Db 199 VVCARPLEGTSFTVDQANRPHCVDPYHKQVAPRCVSEPIPEPGRDETFRVVALDKNF 258
QY 245 HHNCFSCARCSTSLV---GQGFVPDGDQVLCOGCSQA 278
Db 259 HMKYKCEDCGKPLSTEADNNGCFPLDGHVLCRCKHTA 296
```

Search completed: August 22, 2003, 20:09:25
Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:57:47 ; Search time 384 Seconds
(without alignments)
634.636 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSESFDPKCNESLYGRKYI.....QGFPDGDVLCQGCQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata1/paa/PTUS_COMB.pep.*

2: /cgn2_6/ptodata1/paa/US06_COMB.pep.*

3: /cgn2_6/ptodata1/paa/US07_COMB.pep.*

4: /cgn2_6/ptodata1/paa/US080_COMB.pep.*

5: /cgn2_6/ptodata1/paa/US081_COMB.pep.*

6: /cgn2_6/ptodata1/paa/US082_COMB.pep.*

7: /cgn2_6/ptodata1/paa/US083_COMB.pep.*

8: /cgn2_6/ptodata1/paa/US084_COMB.pep.*

9: /cgn2_6/ptodata1/paa/US085_COMB.pep.*

10: /cgn2_6/ptodata1/paa/US086_COMB.pep.*

11: /cgn2_6/ptodata1/paa/US087_COMB.pep.*

12: /cgn2_6/ptodata1/paa/US088_COMB.pep.*

13: /cgn2_6/ptodata1/paa/US089_COMB.pep.*

14: /cgn2_6/ptodata1/paa/US090_COMB.pep.*

15: /cgn2_6/ptodata1/paa/US091_COMB.pep.*

16: /cgn2_6/ptodata1/paa/US092_COMB.pep.*

17: /cgn2_6/ptodata1/paa/US093_COMB.pep.*

18: /cgn2_6/ptodata1/paa/US094_COMB.pep.*

19: /cgn2_6/ptodata1/paa/US095_COMB.pep.*

20: /cgn2_6/ptodata1/paa/US096_COMB.pep.*

21: /cgn2_6/ptodata1/paa/US097A_COMB.pep.*

22: /cgn2_6/ptodata1/paa/US097B_COMB.pep.*

23: /cgn2_6/ptodata1/paa/US098_COMB.pep.*

24: /cgn2_6/ptodata1/paa/US099A_COMB.pep.*

25: /cgn2_6/ptodata1/paa/US099B_COMB.pep.*

26: /cgn2_6/ptodata1/paa/US100_COMB.pep.*

27: /cgn2_6/ptodata1/paa/US101_COMB.pep.*

28: /cgn2_6/ptodata1/paa/US102_COMB.pep.*

29: /cgn2_6/ptodata1/paa/US103_COMB.pep.*

30: /cgn2_6/ptodata1/paa/US104_COMB.pep.*

31: /cgn2_6/ptodata1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1635	100.0	280	22	US-09-786-135-1 Sequence 1, Appl1
2	1635	100.0	280	31	US-60-453-050-9996 Sequence 9996, Ap

3	1635	100.0	280	31	US-60-453-135-9996 Sequence 9996, Ap
4	1635	100.0	280	31	US-60-466-412-9996 Sequence 9996, Ap
5	1635	100.0	287	24	US-09-949-016-9099 Sequence 9099, Ap
6	1625	99.4	280	22	US-09-791-537-141395 Sequence 141395, Ap
7	1615	98.8	280	22	US-09-791-537-141479 Sequence 141479, Ap
8	1595	97.6	280	21	US-09-724-676-62747 Sequence 62747, A
9	1578.5	96.5	280	21	US-09-724-676A-62747 Sequence 62747, A
10	1569	96.0	280	22	US-09-791-537-1413101 Sequence 143101, A
11	1563	95.6	280	22	US-09-791-537-141433 Sequence 141433, A
12	1563	95.6	280	22	US-09-791-537-141432 Sequence 141432, A
13	998	61.0	279	22	US-09-791-537-87691 Sequence 87691, A
14	991	60.6	279	22	US-09-791-537-2014 Sequence 2014, Ap
15	974	59.6	279	21	US-09-724-676-62522 Sequence 62522, A
16	974	59.6	279	21	US-09-724-676-62527 Sequence 62527, A
17	974	59.6	279	21	US-09-724-676-62538 Sequence 62538, A
18	974	59.6	279	21	US-09-724-676A-62522 Sequence 62522, A
19	974	59.6	279	21	US-09-724-676A-62527 Sequence 62527, A
20	974	59.6	279	21	US-09-724-676A-62538 Sequence 62538, A
21	974	59.6	279	22	US-09-791-537-128113 Sequence 128113, Ap
22	974	59.6	279	31	US-60-389-987-169 Sequence 169, App
23	974	59.6	279	31	US-60-412-418-169 Sequence 169, App
24	974	59.6	279	31	US-60-452-680-21669 Sequence 21669, A
25	974	59.6	279	31	US-60-452-680-21670 Sequence 21670, A
26	974	59.6	327	28	US-10-206-021-538 Sequence 538, App
27	968	59.2	279	22	US-09-791-537-100836 Sequence 100836, Ap
28	951.5	58.2	304	21	US-09-724-676-62520 Sequence 62520, A
29	951.5	58.2	304	21	US-09-724-676-62521 Sequence 62521, A
30	951.5	58.2	304	21	US-09-724-676-62525 Sequence 62525, A
31	951.5	58.2	304	21	US-09-724-676-62526 Sequence 62526, A
32	951.5	58.2	304	21	US-09-724-676-62530 Sequence 62530, A
33	951.5	58.2	304	21	US-09-724-676-62531 Sequence 62531, A
34	951.5	58.2	304	21	US-09-724-676A-62520 Sequence 62520, A
35	951.5	58.2	304	21	US-09-724-676A-62521 Sequence 62521, A
36	951.5	58.2	304	21	US-09-724-676A-62526 Sequence 62526, A
37	951.5	58.2	304	21	US-09-724-676A-62530 Sequence 62530, A
38	951.5	58.2	304	21	US-09-724-676A-62531 Sequence 62531, A
39	951.5	58.2	304	22	US-09-791-537-143019 Sequence 143019, Ap
40	922	56.4	525	22	US-09-522-689-1 Sequence 1, Appl1
41	866	53.0	284	19	US-09-522-689A-2 Sequence 2, Appl1
42	866	53.0	284	22	US-09-791-537-141508 Sequence 141508, Ap
43	866	53.0	284	22	US-09-791-537-141507 Sequence 141507, Ap
44	865	52.9	284	22	US-09-949-016-11587 Sequence 11587, A
45	865	52.9	290	24	US-09-949-016-11587 Sequence 11587, A

ALIGNMENTS

RESULT 1

US-09-786-135-1

; Sequence 1, Application US/09786135

; GENERAL INFORMATION:

; APPLICANT: INCYTE PHARMACEUTICALS, INC.

; APPLICANT: TANG, Y. Tom

; APPLICANT: GUEGLER, Karl J.

; APPLICANT: CORLEY, Neil C.

; APPLICANT: PATTERSON, Chandra

; TITLE OF INVENTION: LIM DOMAIN PROTEIN HOMOLOG

; FILE REFERENCE: PF-0585 PCT

; CURRENT APPLICATION NUMBER: US/09/786,135

; CURRENT FILING DATE: 2002-08-27

; PRIOR APPLICATION NUMBER: 09/146,506; unassigned

; PRIOR FILING DATE: 1998-09-03; 1998-09-03

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PERL Program

; SEQ ID NO 1

; LENGTH: 280

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No: 1925017CDI

US-09-786-135-1

Query Match 100.0%; Score 1635; DB 22; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.le-121;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSESFDCAKNESLYGRKYIOTDSGPYCPYDNTFANTCAECQOLIGHDSRELFYEDRH 60
DB 1 MSESFDCAKNESLYGRKYIOTDSGPYCPYDNTFANTCAECQOLIGHDSRELFYEDRH 60
QY 61 FHEGFCRCRCORSRLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
DB 61 FHEGFCRCRCORSRLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFLSCGCEQPLGSRSFVPDKGAHYCPYENKFAPCARCCKTLTQGGVYTRDQ 180
DB 121 TWHEHCFLSCGCEQPLGSRSFVPDKGAHYCPYENKFAPCARCCKTLTQGGVYTRDQ 180
QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPVGLGGKYVSFE 240
DB 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPVGLGGKYVSFE 240
QY 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQCGCSQAGP 280
DB 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQCGCSQAGP 280

RESULT 2

US-60-453-050-9996
; Sequence 9996, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9996
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-9996

Query Match 100.0%; Score 1635; DB 31; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.le-121;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSESFDCAKNESLYGRKYIOTDSGPYCPYDNTFANTCAECQOLIGHDSRELFYEDRH 60
DB 1 MSESFDCAKNESLYGRKYIOTDSGPYCPYDNTFANTCAECQOLIGHDSRELFYEDRH 60
QY 61 FHEGFCRCRCORSRLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
DB 61 FHEGFCRCRCORSRLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFLSCGCEQPLGSRSFVPDKGAHYCPYENKFAPCARCCKTLTQGGVYTRDQ 180
DB 121 TWHEHCFLSCGCEQPLGSRSFVPDKGAHYCPYENKFAPCARCCKTLTQGGVYTRDQ 180
QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPVGLGGKYVSFE 240
DB 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPVGLGGKYVSFE 240
QY 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQCGCSQAGP 280
DB 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQCGCSQAGP 280

RESULT 3

US-60-453-135-9996
; Sequence 9996, Application US/60453135

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9996
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-9996

Query Match 100.0%; Score 1635; DB 31; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.le-121;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSESFDCAKNESLYGRKYIOTDSGPYCPYDNTFANTCAECQOLIGHDSRELFYEDRH 60
DB 1 MSESFDCAKNESLYGRKYIOTDSGPYCPYDNTFANTCAECQOLIGHDSRELFYEDRH 60
QY 61 FHEGFCRCRCORSRLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
DB 61 FHEGFCRCRCORSRLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFLSCGCEQPLGSRSFVPDKGAHYCPYENKFAPCARCCKTLTQGGVYTRDQ 180
DB 121 TWHEHCFLSCGCEQPLGSRSFVPDKGAHYCPYENKFAPCARCCKTLTQGGVYTRDQ 180
QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPVGLGGKYVSFE 240
DB 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPVGLGGKYVSFE 240
QY 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQCGCSQAGP 280
DB 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQCGCSQAGP 280

RESULT 4

US-60-466-412-9996
; Sequence 9996, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9996
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-466-412-9996

Query Match 100.0%; Score 1635; DB 31; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.le-121;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSESFDCAKNESLYGRKYIOTDSGPYCPYDNTFANTCAECQOLIGHDSRELFYEDRH 60
DB 1 MSESFDCAKNESLYGRKYIOTDSGPYCPYDNTFANTCAECQOLIGHDSRELFYEDRH 60
QY 61 FHEGFCRCRCORSRLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
DB 61 FHEGFCRCRCORSRLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFLSCGCEQPLGSRSFVPDKGAHYCPYENKFAPCARCCKTLTQGGVYTRDQ 180

Db 121 TWHEHCFCLSCGCEQPLGSRSFVPDKGAHYCPVCYENKFAFPCARCSKTLTGQGGVYRDQP 180
QY 181 WHRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYYVSFE 240
Db 181 WHRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYYVSFE 240
QY 241 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCQGCQAGP 280
Db 241 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCQGCQAGP 280

RESULT 5

US-09-949-016-9099
; Sequence 9099.. Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9099
; LENGTH: 287
; TYPE: PRP
; ORGANISM: Human
US-09-949-016-9099

Query Match 100.0%; Score 1635; DB 24; Length 287;
Best Local Similarity 100.0%; Pred. No. 2.1e-121;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSEFDCAKCNESLYGRKYYQTDSGPYCPYDNTFANTCAECQOLIGHDSRELFFEDRH 60
Db 8 MSSEFDCAKCNESLYGRKYYQTDSGPYCPYDNTFANTCAECQOLIGHDSRELFFEDRH 67
QY 61 FHGCFRCRCQRLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120
Db 68 FHGCFRCRCQRLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 127
QY 121 TWHEHCFCLSCGCEQPLGSRSFVPDKGAHYCPVCYENKFAFPCARCSKTLTGQGGVYRDQP 180
Db 128 TWHEHCFCLSCGCEQPLGSRSFVPDKGAHYCPVCYENKFAFPCARCSKTLTGQGGVYRDQP 187
QY 181 WHRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYYVSFE 240
Db 188 WHRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYYVSFE 247
QY 241 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCQGCQAGP 280
Db 248 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCQGCQAGP 287

RESULT 6

US-09-791-537-141395
; Sequence 141395.. Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 141395
; LENGTH: 280
; TYPE: PRP
; ORGANISM: Mus musculus
US-09-791-537-141395

Query Match 99.4%; Score 1625; DB 22; Length 280;
Best Local Similarity 99.3%; Pred. No. 1.3e-120;
Matches 278; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSEFDCAKCNESLYGRKYYQTDSGPYCPYDNTFANTCAECQOLIGHDSRELFFEDRH 60
Db 1 MSSEFDCAKCNESLYGRKYYQTDSGPYCPYDNTFANTCAECQOLIGHDSRELFFEDRH 60
QY 61 FHGCFRCRCQRLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120
Db 61 FHGCFRCRCQRLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFCLSCGCEQPLGSRSFVPDKGAHYCPVCYENKFAFPCARCSKTLTGQGGVYRDQP 180
Db 121 TWHEHCFCLSCGCEQPLGSRSFVPDKGAHYCPVCYENKFAFPCARCSKTLTGQGGVYRDQP 180
QY 181 WHRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYYVSFE 240
Db 181 WHRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYYVSFE 240
QY 241 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCQGCQAGP 280
Db 241 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCQGCQAGP 280

RESULT 7

US-09-791-537-141479
; Sequence 141479.. Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 141479
; LENGTH: 280
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-791-537-141479

Query Match 98.8%; Score 1615; DB 22; Length 280;
Best Local Similarity 98.9%; Pred. No. 8.1e-120;
Matches 277; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSSEFDCAKCNESLYGRKYYQTDSGPYCPYDNTFANTCAECQOLIGHDSRELFFEDRH 60
Db 1 MSSEFDCAKCNESLYGRKYYQTDSGPYCPYDNTFANTCAECQOLIGHDSRELFFEDRH 60
QY 61 FHGCFRCRCQRLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120
Db 61 FHGCFRCRCQRLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFCLSCGCEQPLGSRSFVPDKGAHYCPVCYENKFAFPCARCSKTLTGQGGVYRDQP 180
Db 121 TWHEHCFCLSCGCEQPLGSRSFVPDKGAHYCPVCYENKFAFPCARCSKTLTGQGGVYRDQP 180
QY 181 WHRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYYVSFE 240
Db 181 WHRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYYVSFE 240

QY 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280
|||||
Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280

RESULT 8

US-09-724-676-62747
; Sequence 62747, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 62747
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-62747

Query Match 97.6%; Score 1595; DB 21; Length 280;
Best Local Similarity 97.1%; Pred. No. 3.1e-118;
Matches 272; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSESFDCAKNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60
|||||
Db 1 MSESFDCAKNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60
QY 61 FHEGFCRCRCQSLADEPFTCDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
|||||
Db 61 FHEGFCRCRCQSLADEPFTCDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTYRDP 180
|||||
Db 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTYRDP 180
QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVGLGGKYVSFE 240
|||||
Db 181 WHPKCLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVGLGGKYVSFE 240
QY 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280
|||||
Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280

RESULT 9

US-09-724-676A-62747
; Sequence 62747, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 62747
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-62747

Query Match 97.6%; Score 1595; DB 21; Length 280;
Best Local Similarity 97.1%; Pred. No. 3.1e-118;
Matches 272; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSESFDCAKNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60
|||||
Db 1 MSESFDCAKNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60
QY 61 FHEGFCRCRCQSLADEPFTCDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120

Db 61 FHEGFCRCRCQSLADEPFTCDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTYRDP 180
Db 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTYRDP 180
QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVGLGGKYVSFE 240
Db 181 WHPKCLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVGLGGKYVSFE 240
QY 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280
Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280

RESULT 10

US-09-791-537-143101
; Sequence 143101, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biocomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 143101
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-143101

Query Match 96.5%; Score 1578.5; DB 22; Length 289;
Best Local Similarity 93.4%; Pred. No. 6.6e-117;
Matches 270; Conservative 6; Mismatches 4; Indels 9; Gaps 1;

QY 1 MSESFDCAKNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60
|||||
Db 1 MSESFDCAKNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60
QY 61 FHEGFCRCRCQSLADEPFTCDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
|||||
Db 61 FHEGFCRCRCQSLADEPFTCDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTYRDP 180
Db 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTYRDP 180
QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVGLGGKYVSFE 231
Db 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVGLGGKYVSFE 231
QY 232 GGGKYVSFEEDRHHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280
Db 241 GGGKYVSFEEDRHHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 289

RESULT 11

US-09-791-537-141433
; Sequence 141433, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biocomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 141433
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-141433

Query Match 96.0%; Score 1569; DB 22; Length 280;
Best Local Similarity 95.7%; Pred. No. 3.6e-116;
Matches 268; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSESFDCAKCNESLYGRKVIQTDSPGYPVCYVDNTFANTCABEQQLIGHDSRELFYEDRH 60
DB 1 MSESFDCAKCNESLYGRKVIQTDSPGYPVCYVDNTFANTCABEQQLIGHDSRELFYEDRH 60
QY 61 FHGGFRCRCORSLADEPFTQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
DB 61 FHGGFRCRCORSLADEPFTQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFCLGCGEQPLGSRFPDYGKHAHYCVPCYENKFAFPCARCCKTLTQGGVTVRDQP 180
DB 121 TWHEHCFCLGCGEQPLGSRFPDYGKHAHYCVPCYENKFAFPCARCCKTLTQGGVTVRDQP 180
QY 181 WHRECLVCTGCGTQPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKGYVSFE 240
DB 181 WHPKCLVCTGCGTQPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKGYVSFE 240
QY 241 DRHWHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280
DB 241 DRHWHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

RESULT 12

US-09-791-537-141432
; Sequence 141432, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 141432
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-141432

Query Match 95.6%; Score 1563; DB 22; Length 280;
Best Local Similarity 95.4%; Pred. No. 1.1e-115;
Matches 267; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSESFDCAKCNESLYGRKVIQTDSPGYPVCYVDNTFANTCABEQQLIGHDSRELFYEDRH 60
DB 1 MSESFDCAKCNESLYGRKVIQTDSPGYPVCYVDNTFANTCABEQQLIGHDSRELFYEDRH 60
QY 61 FHGGFRCRCORSLADEPFTQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
DB 61 FHGGFRCRCORSLADEPFTQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFCLGCGEQPLGSRFPDYGKHAHYCVPCYENKFAFPCARCCKTLTQGGVTVRDQP 180
DB 121 TWHEHCFCLGCGEQPLGSRFPDYGKHAHYCVPCYENKFAFPCARCCKTLTQGGVTVRDQP 180
QY 181 WHRECLVCTGCGTQPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKGYVSFE 240
DB 181 WHPKCLVCTGCGTQPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKGYVSFE 240

QY 241 DRHWHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280
DB 241 DRHWHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

RESULT 13

US-09-791-537-87691
; Sequence 87691, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 87691
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-87691

Query Match 61.0%; Score 998; DB 22; Length 279;
Best Local Similarity 55.6%; Pred. No. 8.1e-71;
Matches 154; Conservative 55; Mismatches 68; Indels 0; Gaps 0;

QY 1 MSESFDCAKCNESLYGRKVIQTDSPGYPVCYVDNTFANTCABEQQLIGHDSRELFYEDRH 60
DB 1 MTERFDCCHNCSLYGKVIKKEENPHCVACFEELYANTCEGCTPGCDCKDLSYKDRH 60
QY 61 FHGGFRCRCORSLADEPFTQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
DB 61 FHGGFRCRCORSLADEPFTQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
QY 121 TWHEHCFCLGCGEQPLGSRFPDYGKHAHYCVPCYENKFAFPCARCCKTLTQGGVTVRDQP 180
DB 121 TWHEHCFCLGCGEQPLGSRFPDYGKHAHYCVPCYENKFAFPCARCCKTLTQGGVTVRDQP 180
QY 181 WHRECLVCTGCGTQPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKGYVSFE 240
DB 181 WHRECLVCTGCGTQPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKGYVSFE 240
QY 241 DRHWHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQ 277
DB 241 ERQWHDNCFNCKKCSLVGRGFLTERDDILCPDCGK 277

RESULT 14

US-09-791-537-2014
; Sequence 2014, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2014
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-2014

Query Match 60.6%; Score 991; DB 22; Length 279;
Best Local Similarity 54.9%; Pred. No. 2.9e-70;
Matches 152; Conservative 57; Mismatches 68; Indels 0; Gaps 0;

```

Qy 1 MSESFDCAKNESLYGRKYIQTDSGPCYPCYDNTFANTCAEQOOLIGHDSRELFYEDRH 60
Db 1 MTERFDCHHCNESLYGKYLKEENPHCVACFEELYANTCECGTPIGCDCKDLSYKDRH 60
Qy 61 FHEGCFRCRCQRSLADEPFTCQDSSELLCNDYCSAFSSQCSAGETVMPGSRKLEYGGQ 120
Db 61 WHEGCFHCSRGSSLVDPFAAKBEQLLCTDCYSNEYSSKQCECKTITMPGTRMEYKGS 120
Qy 121 TWHEHCFCLSCGEOPLGSRSPVDPKGAHYCPVCYENKFAPCARCSKTLTGGVYTRDOP 180
Db 121 SWHETCFQCRCQPIGTGKSFIPKFNQNFVPCYKQYALQCVQCKPITTTGGVYTRDOP 180
Qy 181 WHRECLVTCQOTPLAGQQFTSRDEDPYCVACFGELFAPKSCSKRPVIGLGGKGYVSFE 240
Db 181 WHKECFVCTACKKQLSGORFTARDEFFVCLTCFDLYAKKACAGCTNPISGLGGTKYISFE 240
Qy 241 DRWHHNCFCARCSSTLVGGVFDGQVLCQGCSCQ 277
Db 241 EROMHNDCFNCKGCSLVGRGFLTERDDILCPDCGK 277

```

RESULT 15

```

US-09-724-676-62522
; Sequence 62522, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 62522
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-62522

```

```

Query Match          59.6%; Score 974; DB 21; Length 279;
Best Local Similarity 53.8%; Pred. No. 6.5e-69;
Matches 149; Conservative 59; Mismatches 69; Indels 0; Gaps 0;

Qy 1 MSESFDCAKNESLYGRKYIQTDSGPCYPCYDNTFANTCAEQOOLIGHDSRELFYEDRH 60
Db 1 MTERFDCHHCNESLYGKYLKEENPHCVACFEELYANTCECGTPIGCDCKDLSYKDRH 60
Qy 61 FHEGCFRCRCQRSLADEPFTCQDSSELLCNDYCSAFSSQCSAGETVMPGSRKLEYGGQ 120
Db 61 WHEGCFHCSRGSSLVDPFAAKBEQLLCTDCYSNEYSSKQCECKTITMPGTRMEYKGS 120
Qy 121 TWHEHCFCLSCGEOPLGSRSPVDPKGAHYCPVCYENKFAPCARCSKTLTGGVYTRDOP 180
Db 121 SWHETCFQCRCQPIGTGKSFIPKFNQNFVPCYKQYALQCVQCKPITTTGGVYTRDOP 180
Qy 181 WHRECLVTCQOTPLAGQQFTSRDEDPYCVACFGELFAPKSCSKRPVIGLGGKGYVSFE 240
Db 181 WHKECFVCTACKKQLSGORFTARDEFFVCLTCFDLYAKKACAGCTNPISGLGGTKYISFE 240
Qy 241 DRWHHNCFCARCSSTLVGGVFDGQVLCQGCSCQ 277
Db 241 EROMHNDCFNCKGCSLVGRGFLTERDDILCPDCGK 277

```

Search completed: August 22, 2003, 20:07:49
Job time : 386 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:58:12 ; Search time 24 Seconds
(without alignments)
313.637 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSESFDCAKNESLYGRKYI.....QGFVPDQVLCQGCSCQACP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 128601 seqs, 26883176 residues

Total number of hits satisfying chosen parameters: 128601

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	974	59.6	279	6	US-10-408-765A-169
2	974	59.6	279	7	US-60-490-890-1761
3	437.5	26.8	150	6	US-10-408-765A-210
4	305.5	18.7	264	7	US-60-487-610-1707
5	305.5	18.7	461	6	US-10-293-244-1620
6	305.5	18.7	486	6	US-10-293-244-3586
7	303.5	18.6	734	6	US-10-408-765A-324
8	292.5	17.9	296	6	US-10-374-979-102
9	275	16.8	831	6	US-10-408-765A-2503
10	265	16.2	615	6	US-10-293-244-1901
11	265	16.2	615	6	US-10-408-765A-913
12	262	16.0	539	6	US-10-408-765A-1767
13	260	15.9	212	6	US-10-275-595A-21
14	257.5	15.7	572	6	US-10-374-979-107
15	245.5	15.0	645	6	US-10-293-244-3869
16	228	13.9	431	6	US-10-374-979-90
17	220	13.5	412	7	US-60-490-890-379
18	220	13.5	421	6	US-10-408-765A-1032
19	220	13.5	421	6	US-10-286-897-2375
20	220	13.5	421	6	US-10-258-898A-2375
21	220	13.5	421	7	US-60-490-890-377
22	220	13.5	437	6	US-10-286-897-5947
23	220	13.5	437	6	US-10-258-898A-5947
24	209	12.8	676	6	US-10-374-979-101
25	200	12.2	194	6	US-10-291-172-224
26	196.5	12.0	365	6	US-10-286-897-3084

ALIGNMENTS

RESULT 1

US-10-408-765A-169

; Sequence 169, Application US/10408765A

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 169

; LENGTH: 279

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-408-765A-169

Query Match 59.6%; Score 974; DB 6; Length 279;

Best Local Similarity 53.8%; Pred. No. 4.le-80;

Matches 149; Conservative 59; Mismatches 69; Indels 0; Gaps 0;

Qy 1 MSESFDCAKNESLYGRKYIQTDSGPCYVPCYDNTFANTCAECQOLIGHDSRELFEYDRH 60

Db 1 MTERFDCCHCNESLFGKKYILRESPYCVVCFETLANTCECGKPIGCDCKLSYKDRH 60

Qy 61 FHEGCFCCRCORSLADEPTCQDSLELLCNDYCASFSSQCSAGETVMPGSRKLEYGGQ 120

Db 61 WHEACFHCSCRNLSLVDPKFAAKEDQLCTDYSNEYSSKQCECKTKIMPGRKMEYKGS 120

Qy 121 TWHEHCFCLSCCEOPGLGSRFPDKGAHYCPVCYENKFAPCARCSKTLTQGGVYTRQP 180

Db 121 SWHETCFICHRCCQPIGTFSKFPKDNQNFVPCYCKEQAHCQVCKKPIITGGVYTRQP 180

Qy 181 WHRELYCTGQCTPLAQOQFTSRDEDPYCVACFELFAPKCSSCKRPITVGLGGKYSVFE 240

Db 181 WKKECFVCTACRKQLSGQRTFARDFAVLCNFCDLVAKKACAGCTNPISGLGKTKYSFE 240

Qy 241 DRWHHNCFCARCSSTSLVCGQFVDPDGVLCQGCSCQ 277

Db 241 ERQWHDNCFNCKCSLSLVGRGFLTERDDILCPDQCK 277

Sequence 3084, Ap

Sequence 6656, Ap

Sequence 6856, Ap

Sequence 1739, Ap

Sequence 600, App

Sequence 404, App

Sequence 1124, Ap

Sequence 1054, Ap

Sequence 3022, Ap

Sequence 3929, Ap

Sequence 1345, Ap

Sequence 2081, Ap

Sequence 1961, Ap

Sequence 2811, Ap

Sequence 528, App

Sequence 1034, Ap

Sequence 3002, Ap


```

; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1620
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-1620

Query Match      18.7%; Score 305.5; DB 6; Length 461;
Best Local Similarity 31.4%; Pred. No. 4.6e-20;
Matches 59; Conservative 24; Mismatches 90; Indels 15; Gaps 3;

QY 101 CSACGETVMPGSRKLEYGGQTWHEHCFCLSCGCEQPLGSRFSFVDPDKGAHYCVPCYENKFP 160
Db 228 CGSCNKPI--AGQVVTALGRAWHPHFVCGGCGTALGGSSFFFEKDGAPFCPECYFERFSP 285
QY 161 RCARCSKTLTQGGVTVRDPQWHRRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPK 220
Db 286 RCQFCNQPIRHKVMVTALGTHWHPEHFCCVSCGPFDEGFEHREGRPYCRDRDLQLFAPR 345
QY 221 CSSCKRPVGLGGKGVSVFEDRHHNCFSCARCSSTSLVGQGFVDPDQVLCQG----- 274
Db 346 CQCGQGPIL----DNVISALSALWHPDCFCVRCRCFAPFGSGGSFFEHEGRPLCENHFHARR 401
QY 275 ---CSQAG 279
Db 402 GSLCATCG 409

RESULT 6
US-10-293-244-3588
; Sequence 3588, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 3588
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```

US-10-293-244-3588

Query Match      18.7%; Score 305.5; DB 6; Length 486;
Best Local Similarity 31.4%; Pred. No. 4.9e-20;
Matches 59; Conservative 24; Mismatches 90; Indels 15; Gaps 3;

QY 101 CSACGETVMPGSRKLEYGGQTWHEHCFCLSCGCEQPLGSRFSFVDPDKGAHYCVPCYENKFP 160
Db 253 CGSCNKPI--AGQVVTALGRAWHPHFVCGGCGTALGGSSFFFEKDGAPFCPECYFERFSP 310
QY 161 RCARCSKTLTQGGVTVRDPQWHRRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPK 220
Db 311 RCQFCNQPIRHKVMVTALGTHWHPEHFCCVSCGPFDEGFEHREGRPYCRDRDLQLFAPR 370
QY 221 CSSCKRPVGLGGKGVSVFEDRHHNCFSCARCSSTSLVGQGFVDPDQVLCQG----- 274
Db 371 CQCGQGPIL----DNVISALSALWHPDCFCVRCRCFAPFGSGGSFFEHEGRPLCENHFHARR 426
QY 275 ---CSQAG 279
Db 427 GSLCATCG 434

RESULT 7
US-10-408-765A-324
; Sequence 324, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 324
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-324

Query Match      18.6%; Score 303.5; DB 6; Length 734;
Best Local Similarity 36.3%; Pred. No. 1.1e-19;
Matches 65; Conservative 21; Mismatches 86; Indels 7; Gaps 5;

QY 98 SSQCSACG--ETVMPGSRKLEYGGQTWHEHCFCLSCGCEQPLGSRFSFVDPDKGAHYCVPCYE 155
Db 552 SSRTPLCGHCNNVIRGP-FLVAMGRSWHPDEFTCAICTSLADVCFVEEQNNVYCERCYE 610
QY 156 NKFAPCARCASKTLTQGGVTVR-DQWHRRECLVCTGCQTPLAGQQFTSRDEDPYCVACFG 214
Db 611 QFFAPLCAKCN-TKINGEVMHALRQTHWTTTCFVCAACKKPGNSLFHMEDEGPYCDYI 669
QY 215 ELFAPKCSSCKRPVGLGGKGVSVFEDRHHNCFSCARCSSTSLVGQGFVDPDQVLCQ 273
Db 670 NLFSTKCHGCDPVE--AGDKFIALGHTWHTDTCFICAVCHVNLEGQPFYSSKDRPLCK 726

RESULT 8
US-10-374-979-102
; Sequence 102, Application US/10374979
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04

```

```

RESULT 9
US-10-408-765A-2503
; Sequence 2503, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fany, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2503
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2503

Query Match          16.8%; Score 275; DB 6; Length 831;
Best Local Similarity 27.0%; Pred No. 4.6e-17;
Matches 76; Conservative 43; Mismatches 106; Indels 56; Gaps 13;

QY 28 CVPYDNTFANTCAEQQLIGHDSRELFYEDRHFHEGCFRCRCQRSLADEPFTCDQSEL 87
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 51 CLPEEKVPVNSPGE-----KHKIKLLQY-LPHDNEVRYC---QSLSEE-----EKKEL 97

QY 88 LCNDCTCSAFPSQ-----CSAC-----GTVMPGSKRYGG 119
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 98 -----QVFSQKKKEALGRGTYIKLLSRVHMVAVCEQCGLKINGGVAVFASR--AGPG 148

```

; Sequence 913, Application US/10408765A

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 913

; LENGTH: 615

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-913

Query Match 16.2%; Score 265; DB 6; Length 615;

Best Local Similarity 29.9%; Pred. No. 2.4e-16;

Matches 55; Conservative 31; Mismatches 88; Indels 10; Gaps 6;

QY 101 CSAGETVMPGSRKLEYG---GQTWHEHCFLCSCGCEQPLGSRSFVPDKGAHYCVPCYEN 156

DB 186 CRECKQIGGGDIANFASRAGLACWHPQCFCTTCQELLVDLIIFYHVGKVCYGRHHAE 245

QY 157 KFAPCARCSTLTGGVTYRD-QPWHRECLVCTGCQTPLAGQOFTSRDEDPYCVACFGE 215

DB 246 CLRPRCAQCEIIFSPCTEAEGRHWMDFCCFECEASLGORYVMQRSPHCCACYEA 305

QY 216 LFAPKSSCKRPVGLGGKYVSFEDRHH--HNCFSARCSTSLVGQGFVDPGDQVLC- 272

DB 306 RHAECYCDGGEHI-GLDQSQ-MAYEGQHWASDRFCSCRCRALLIGRFLPRRLIFCS 363

QY 273 QGCS 276

DB 364 RACS 367

RESULT 12

US-10-408-765A-1767

; Sequence 1767, Application US/10408765A

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1767

; LENGTH: 539

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-1767

Query Match

Best Local Similarity 16.0%; Score 262; DB 6; Length 539;

Matches 67; Conservative 37; Mismatches 100; Indels 42; Gaps 10;

QY 57 EDRHFHEGFCRCRCQRSIADPFTCCQDSEILCNDICYSAFSSQCSAC-----GEFTVMPG 111

DB :::: ||| | ||| | : | | : | | : | | : | | |

DB 6 ODKYFIHKCFVCKACGCDLAEGGFFVROGEYICTLDYQRYLGTGTRCFSCDQFIERGEVV--- 62

QY 112 SRKLEYGGOTWHEHCFLCSCGCEQPLGSRSFVPDKGAH-YCVPC-----YENKFAP 160

DB 63 ----SALGKTYHPDCFCVCAVCRLPFPFGDRVTFNGKRCMCQKCSLPSVSGSSAHLQGLR 118

QY 161 RCARCSKTLTQG-GVTYRDQPHRECLVCTGCQTPLAGQOFTSRDEDPYCVACFGE 219

DB 119 SCGGCGTEIKNGOALVALDKHMLGCFCKKSC-GKLLNAEYISKDGLPYCEADYHAKFGI 177

QY 220 KSSCKRPVGLGGKYVSFEDRHHNCFSCARCSTSLVGQGFVDPGDQVLCQ----- 274

DB 178 RCDSCKEYIT----GRVLEAGEKHYPSCALCVRC-----GOMFA-EGEBMYLQGGSIWH 227

QY 275 --CSQA 278

DB 228 PACRQA 233

RESULT 13

US-10-275-595A-21

; Sequence 21, Application US/10275595A

; GENERAL INFORMATION:

; APPLICANT: YUE, Henry

; APPLICANT: TANG, Y. Tom

; APPLICANT: AU-YOUNG, Janice

; APPLICANT: LU, Dyung Aina M.

; APPLICANT: BAUGHN, Mariah R. L.

; APPLICANT: HILLMAN, Jennifer L.

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: LAL, Preeti

; APPLICANT: YAO, Monique G.

; APPLICANT: BANDMAN, Olga

; APPLICANT: BURFORD, Neil

; APPLICANT: BATRA, Sajeev

; APPLICANT: KEARNEY, Liam

; APPLICANT: POLICKY, Jennifer L.

; TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS

; FILE REFERENCE: PF-0772 USN

; CURRENT APPLICATION NUMBER: US/10/275,595A

; CURRENT FILING DATE: 2003-06-13

; PRIOR APPLICATION NUMBER: US 60/201,960

; PRIOR FILING DATE: 2000-05-05

; PRIOR APPLICATION NUMBER: US 60/202,729

; PRIOR FILING DATE: 2000-05-08

; PRIOR APPLICATION NUMBER: US 60/209,705

; PRIOR FILING DATE: 2000-06-05

; PRIOR APPLICATION NUMBER: US 60/210,149

; PRIOR FILING DATE: 2000-06-07

; PRIOR APPLICATION NUMBER: US 60/213,215

; PRIOR FILING DATE: 2000-06-21

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: PERL Program

; SEQ ID NO 21

; LENGTH: 212

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No: 2913976CDI

US-10-275-595A-21

Query Match 15.9%; Score 260; DB 6; Length 212;

Best Local Similarity 29.5%; Pred. No. 2.4e-16;

Matches 56; Conservative 26; Mismatches 96; Indels 12; Gaps 5;

QY 101 CSAGETVMPGSRKLEYGGOTWHEHCFLCSCGCEQPLGSRSFVPDKGAHYCVPCYENK 160

DB 22 CAFCHTVPFRELAVEMKRQYHAQCFTCTCRQLAGQSFYKQDGRPLCEPCYQDTL-E 80

QY 161 RCARCSKTLTQGGVTYRDQPHRECLVCTGCQTPLAGQOFT-SRDEDPYCVACFGE 219

DB 81 RCGKCGVVRDHIIRALGOAFHPSCFTCVTCARCIGDESFALGSONEVYCLDDFYRK 140

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:54:46 ; Search time 40 Seconds

(without alignments)
673.181 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSESFDCAKNESLYGRKYI.....QGFVPDQVLCQGCQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1569	96.0	280	2 T09504	LIM-protein FHL3,
2	968	59.2	279	2 JC6565	four-and-a-half LI
3	922	56.4	525	2 T21357	hypothetical prote
4	844	51.6	284	2 JC7686	activator of camp-
5	755.5	46.2	280	2 G01884	LIM protein FHL-1,
6	730	44.6	454	2 T21358	hypothetical prote
7	728.5	44.6	280	2 G02741	skeletal muscle LI
8	718	43.9	279	2 JG0164	LIM protein, FHL4
9	311.5	19.1	444	2 A55071	hydrogen peroxide-
10	306.5	18.7	314	2 JC2324	LIM protein - huma
11	304.5	18.6	557	2 A55933	paxillin - human
12	303.5	18.6	559	2 B55933	paxillin - chicken
13	300.5	18.4	256	2 E88469	protein C28H8.6 [i
14	297.5	18.2	348	2 T16076	LIM protein - Caen
15	293	17.9	455	2 A55050	enigma - human
16	286	17.5	1333	2 T19157	probable metal bin
17	286	17.5	1424	2 T19156	probable metal bin
18	281.5	17.2	575	2 C88346	protein F42G4.3a [
19	281.5	17.2	603	2 T22111	hypothetical prote
20	277.5	17.0	465	2 T29257	hypothetical prote
21	265.5	16.2	176	2 T22110	hypothetical prote
22	259	15.8	200	2 T42678	hypothetical prote
23	257.5	15.7	572	2 G02845	zyxin - human
24	254	15.5	542	2 A44358	zyxin - chicken
25	241	14.7	438	2 T40509	lim domain protein
26	237	14.5	329	2 T20546	hypothetical prote
27	226	13.8	423	2 I48842	testin - mouse
28	212	13.0	192	2 T26544	hypothetical prote
29	195	11.9	194	2 S41761	cysteine-rich prot

beta-cysteine-rich
murine muscle LIM
muscle LIM protein
hypothetical prote
probable actin-bin
hypothetical prote
cysteine-rich prot
finger protein (cl
laminin alpha-1 ch
probable testin DK
cysteine-rich prot
LIM-domain protein
cysteine-rich prot
LIM protein kinase
finger protein (cl
LIM protein - yea

ALIGNMENTS

RESULT 1

T09504

LIM-protein FHL3, skeletal muscle - human

C;Species: Homo sapiens (man)

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Nov-2000

C;Accession: T09504; PD0036

R;Morgan, M.J.; Madgwick, A.J.A.

Biochem. Biophys. Res. Commun. 255, 245-250, 1999

A;Title: The LIM proteins FHL1 and FHL3 are expressed differently in skeletal muscle.

A;Reference number: PD0036; MUID:99160848; PMID:10049693

A;Accession: T09504

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-280 <MOR>

A;Cross-references: EMBL:U60116; NID:g416529; PIDN:AAC04466.2; PID:g4416530

A;Accession: PD0036

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-39 <MO2>

A;Cross-references: GB:U60116

C;Genetics:

A;Map position: 1

C;Keywords: metal binding; zinc finger

Query Match 96.0%; Score 1569; DB 2; Length 280;

Best Local Similarity 95.7%; Pred. No. 6.9e-106;

Matches 268; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MSESFDCAKNESLYGRKYIQTDSGPCYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60

Db 1 MSESFDCAKNESLYGRKYIQTDSGPCYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60

Qy 61 FHEGFCRCRCQRSLADEPFTCCQSELNDCYCSAFSSQCSAGETVMPGSRKLEYGGQ 120

Db 61 FHEGFCRCRCQRSLADEPFTCCQSELNDCYCSAFSSQCSAGETVMPGSRKLEYGGQ 120

Qy 121 TWHEHCFCLSCGCPGLGSRFVPDKGAHYCPVCYENKFAPCARCSTLTGGVYTRDQP 180

Db 121 TWHEHCFCLSCGCPGLGSRFVPDKGAHYCPVCYENKFAPCARCSTLTGGVYTRDQP 180

Qy 181 WHRECLVCTGCTPLAGQQFTSRDEDPYCVACFELFAPKSCSKRPITVLGGGKYVSFE 240

Db 181 WHRECLVCTGCTPLAGQQFTSRDEDPYCVACFELFAPKSCSKRPITVLGGGKYVSFE 240

Qy 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Qy 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Qy 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Qy 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Qy 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Qy 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Qy 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280

C:Species: Homo sapiens (man)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jul-2000
C:Accession: J05565
R:Chan, K.K.; Tsui, S.K.W.; Lee, S.M.Y.; Luk, S.C.W.; Liew, C.C.; Fung, K.P.; Waye, M.M.
Gene 210, 345-350, 1998
A:Title: Molecular cloning and characterization of FHL2, a novel LIM domain protein preferentially expressed in muscle
A:Reference number: J05565; MUID:98248917; PMID:9573400
A:Accession: J05565
A:Molecule type: mRNA
A:Residues: 1-279 <CHA>
A:Cross-references: GB:U29332; NID:q1845201; PIDN:AAC52073.1; PID:q1377897
A:Experimental source: heart
C:Genetics:
A:Map position: 2q12-2q13
C:Superfamily: LIM metal-binding repeat homology
C:Keywords: cardiac muscle; heart; zinc finger
F:221-275/Domain: LIM metal-binding repeat homology <LIM>

Query Match 59.2%; Score 968; DB 2; Length 279;
Best Local Similarity 53.4%; Pred. No. 1e-62;
Matches 148; Conservative 59; Mismatches 70; Indels 0; Gaps 0;

QY 1 MSEPDCACNLSYGRKYIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRELFEYDRH 60
DB 1 MTERDCHHCNLSLFGKYLILREESPYCVVCFETLFTANTCEGKPIGDCDKLSTKDRH 60

QY 61 PHEGFCRCRCQRLADEPFTQDSSELLCNDYCYSFSSQCSACGETVMPGSRKLEYGGQ 120
DB 61 WHEACFHCSCQRLSLVDKPAFAKEDQLLDCYSNEYSKQCECKTTPGTRKMEYKGS 120

QY 121 TWHEHCFCLSGCEQPLGSRFVDPKGAHYCVPCYENKFAPRCARCKSKTLTQGGVYRDQP 180
DB 121 SWHETCFICRCCQPIGTGKSTFKPNQNFVPCYKQHAMQCVCKMPITTTGGVYREQP 180

QY 181 WRECLVCTGCGTLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKVVSPDE 240
DB 181 WKECFVCTACKQKLSGQRTARDPAYCLNCGFTARDYAKKACAGTNPISGLGGTRYISPE 240

QY 241 DRWHHNCFCARCSTSLVGQGFVDPGQVLCQCGSQ 277
DB 241 ERQWHDNCFNCKSLSLVGRGFLTERDILCPDCGK 277

RESULT 3
T21357
hypothetical protein F25H5.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21357
R:Steward, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19412
A:Accession: T21357
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-525 <WIL>
A:Cross-references: EMBL:281068; PIDN:CAB02980.1; GSPDB:GN00019; CESP:F25H5.1a
A:Experimental source: clone F25H5
C:Genetics:
A:Gene: CESP:F25H5.1a
A:Map position: 1
A:Introns: 35/3; 114/3; 153/3; 186/3; 253/3; 288/1; 312/1; 431/1; 478/3

Query Match 56.4%; Score 922; DB 2; Length 525;
Best Local Similarity 50.4%; Pred. No. 3.4e-59;
Matches 139; Conservative 64; Mismatches 73; Indels 0; Gaps 0;

QY 2 SESEPCACNLSYGRKYIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRELFEYDRH 61
DB 203 SDHFCWCQDQTLTGQYIMRDEQPYCKYEDVFANQDCEAKPIGIDSLSTYKDKHW 262

QY 62 HEGCFRCRCRCQRLADEPFTQDSSELLCNDYCYSFSSQCSACGETVMPGSRKLEYGGQ 121
DB 62 WKECFVCTACKQKLSGQRTARDPAYCLNCGFTARDYAKKACAGTNPISGLGGTRYISPE 240

DB 263 HEHCFCLSMCKISLVDMPFGSKNDRIFCSNCDYDQAFATRCDCGNEIFRAGMKMEYKKGQ 322
QY 122 WHEHCFCLSGCEQPLGSRFVDPKGAHYCVPCYENKFAPRCARCKSKTLTQGGVYRDQPW 181
DB 323 WIDKFCFCAHCKLAIGTKSFIPKNDVDFCGPYEKEFATRCCKKVIITAGGVTYKNEPW 382
QY 182 HRECLVCTGCGTLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKVVSPDE 241
DB 383 HRECFCCITNCNSSLAGQRFSTDEKPYCANCYGLDFAKRCNACTPITGIGGAKFISFED 442

QY 242 RHWHHNCFCARCSTSLVGQGFVDPGQVLCQCGSQ 277
DB 443 RHWHNDCFCAOCTTSLVGKGFITDGHGHEILCPECAK 478

RESULT 4
JC7686
activator of cAMP-responsive element modulator, testis - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: JC7686
R:Palermo, I.; Litrico, L.; Emmanuele, G.; Giuffrida, V.; Sassone-Corsi, P.; De Cesar
Biochem. Biophys. Res. Commun. 283, 406-411, 2001
A:Title: Cloning and expression of activator of CREM in testis in human testicular ti
A:Reference number: JC7686; MUID:21226154; PMID:11327716
A:Contents: Testis
A:Accession: JC7686
A:Molecule type: mRNA
A:Residues: 1-284 <PAL>
A:Cross-references: GB:AF278541
C:Comment: This protein activates cAMP-responsive element modulator in a phosphorylat
of human spermatogenesis.
C:Genetics:
A:Gene: act
A:Map position: 6q
A:Introns: 53/3; 112/3; 168/3; 231/3
C:Keywords: spermatogenesis; testis

Query Match 51.6%; Score 844; DB 2; Length 284;
Best Local Similarity 47.8%; Pred. No. 8.6e-54;
Matches 130; Conservative 51; Mismatches 91; Indels 0; Gaps 0;

QY 5 FDCACNLSYGRKYIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRELFEYDRHFG 64
DB 6 FYCQYCTASLLGKLYLVKVDSPYCVTCYDRVSNYCECKKPTIESDKDFCYKRVHHEG 65

QY 65 CFRCRCRCQRLADEPFTQDSSELLCNDYCYSFSSQCSACGETVMPGSRKLEYGGQWHE 124
DB 66 CFKCTKCNHSLVEKFPAAKDERLLCTECYSNCSCKCFHCKRTIMPGRKMEFKGNYWHE 125

QY 125 HCFLSGCEQPLGSRFVDPKGAHYCVPCYENKFAPRCARCKSKTLTQGGVYRDQPWHR 184
DB 126 TCFVCNCRQPIGTGKPLISKESGNCVPCFEKEFAHCNFKKVTSGGTTFCQDLWHE 185

QY 185 CLVCTGCGTLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKVVSPEDRW 244
DB 186 CFLCSGCRKDLCEQFMSRDDYFPCMDCYNHLVANKVACSKPISGLTCAKFKICFQDSQW 245

QY 245 HHNCFSCARCSTSLVGQGFVDPGQVLCQCGCS 276
DB 246 HSECFNCRGCSVSLVGKGFITQNKEIFCQKCS 277

RESULT 5
G01884
LIM protein FHL-1, cardiac - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Jul-1999
C:Accession: G01884
R:Waye, M.M.Y.
submitted to the EMBL Data Library, June 1995
A:Reference number: H00697
A:Accession: G01884

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-280 <SWA>
 A:Cross-references: EMBL:U29538; NID:g1377898; PID:g1377899
 C:Genetics:
 A:Gene: FHL-1

C:Superfamily: LIM metal-binding repeat homology
 C:Keywords: cardiac muscle; duplication; heart
 F:40-92/Domain: LIM metal-binding repeat homology <LIM1>
 F:101-153/Domain: LIM metal-binding repeat homology <LIM2>
 F:162-212/Domain: LIM metal-binding repeat homology <LIM3>
 F:221-276/Domain: LIM metal-binding repeat homology <LIM4>

Query Match 46.2%; Score 755.5; DB 2; Length 280;
 Best Local Similarity 44.2%; Pred. No. 1.9e-47;
 Matches 123; Conservative 50; Mismatches 104; Indels 1; Gaps 1;

QY 1 MSESFDCAKCNESLYGRKVIQTDSPGYCPVPCYDNTFANTCAEQQLIGHDSRELFPYEDRH 60
 DB 1 MAEKFDCHYCRVPLQGGKYVQKHCCCLCKDFCANTVCBCRPIGADSKEVHYKNRF 60
 QY 61 FHGFCRCRCORSRLADPEFTQDSSELLCNDVCYCSAFSSQCSACGETVMPGSRKLEYGQ 120
 DB 61 WHDTFCRCACLUPLANETFAKDKILCNKTTREDSSCKGCFKAI VAGDQNVYKGT 120
 QY 121 TWHEHCFCLSGCGEPLGSRSFVPDKGAHYCPYENKFAPRCARCSKTLTQGGVYRDQP 180
 DB 121 VWHKDCFTCSNCKQVIGTSGFPFGEDFYCVTCHETKFAKHCVCNKAITSGGITYQDQP 180
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLG-GGKYVSF 239
 DB 181 WHADCFVCTCSKKLGAFAFTVEDQYCYVDFKFAKCKAGCNKPNITGFGKGSVVAY 240
 QY 240 EDRHWHNCFSCARCSSTLVGGFVDPDGVLCQCSQ 277
 DB 241 EGQSHDYCFHCKKCSVNLANKRFVHQEQVYCPDCAK 278

RESULT 6

T21358
 hypothetical protein F25H5.1b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21358
 R:Steward, C.

A:Reference number: Z19412
 submitted to the EMBL Data Library, October 1996

A:Accession: T21358

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-454 <SWIL>

A:Cross-references: EMBL:281068; PIDN:CAB02981.1; GSPDB:GN00019; CESP:F25H5.1b

A:Experimental source: clone F25H5

C:Genetics:

A:Gene: CESP:F25H5.1b

A:Map position: 1

A:Introns: 35/3; 114/3; 153/3; 186/3; 253/3; 288/1; 312/1

Query Match 44.6%; Score 730; DB 2; Length 454;
 Best Local Similarity 48.5%; Pred. No. 1.9e-45;
 Matches 111; Conservative 51; Mismatches 67; Indels 0; Gaps 0;

QY 2 SSESFDCAKCNESLYGRKVIQTDSPGYCPVPCYDNTFANTCAEQQLIGHDSRELFPYEDRH 61
 DB 203 SHFCFCWQCDQTLTGRIYMRDEQPYCIKYEDVFANQCDCAKPIGIDSKDLSYDKRW 262
 QY 62 HEHCFRCRCORSRLADPEFTQDSSELLCNDVCYCSAFSSQCSACGETVMPGSRKLEYGQ 121
 DB 263 HEHCFCLSNCKISLVDMPSGKNDRIFCNSCYDQAFATRCDCNCFRAGMKMEYKQK 322
 QY 122 WHRECLVCTGCGEPLGSRSFVPDKGAHYCPYENKFAPRCARCSKTLTQGGVYRDQP 181
 DB 323 WHDKFCFCAHCKLAIGTKSFIPKNDVFCGPGCYEKFATRCCKKVIITAGGVYTKNEPW 382

QY 182 HRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIG 230
 DB 383 HRECFCCINCSSLAGQRTSKDEPYCANCYGDLFKACNACTKPTIG 431

RESULT 7

G02741

skeletal muscle LIM-protein SLIM1 - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Jul-1999

C:Accession: JG4893; G02741

R:Morgan, M.J.; Madgwick, A.J.A.

Biochem. Biophys. Res. Commun. 225, 632-638, 1996

A:Title: SLIM defines a novel family of LIM-proteins expressed in skeletal muscle.

A:Reference number: JG4893; MUID:96354835; PMID:8753811

A:Accession: JG4893

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-280 <MO2>

A:Cross-references: EMBL:U60115; NID:g1381807; PID:g1381808

C:Superfamily: LIM metal-binding repeat homology

F:162-212/Domain: LIM metal-binding repeat homology <LIM>

Query Match 44.6%; Score 728.5; DB 2; Length 280;
 Best Local Similarity 42.8%; Pred. No. 1.7e-45;
 Matches 119; Conservative 50; Mismatches 108; Indels 1; Gaps 1;

QY 1 MSESFDCAKCNESLYGRKVIQTDSPGYCPVPCYDNTFANTCAEQQLIGHDSRELFPYEDRH 60
 DB 1 MAEKFDCHYCRVPLQGGKYVQKHCCCLCKDFCANTVCBCRPIGADSKEVHYKNRF 60
 QY 61 FHGFCRCRCORSRLADPEFTQDSSELLCNDVCYCSAFSSQCSACGETVMPGSRKLEYGQ 120
 DB 61 WHDTFCRCACLUPLANETFAKDKILCNKTTREDSSCKGCFKAI VAGDQNVYKGT 120
 QY 121 TWHEHCFCLSGCGEPLGSRSFVPDKGAHYCPYENKFAPRCARCSKTLTQGGVYRDQP 180
 DB 121 VWHKDCFTCSNCKQVIGTSGFPFGEDFYCVTCHETKFAKHCVCNKAITSGGITYQDQP 180
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLG-GGKYVSF 239
 DB 181 WHADCFVCTCSKKLGAFAFTVEDQYCYVDFKFAKCKAGCNKPNITGFGKGSVVAY 240

RESULT 8

JG0164

LIM protein, FHL4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000

C:Accession: JG0164

R:Morgan, M.J.; Madgwick, A.J.A.

Biochem. Biophys. Res. Commun. 255, 251-255, 1999

A:Title: The fourth member of the FHL family of LIM proteins is expressed exclusively

A:Reference number: JG0164; MUID:99160849; PMID:10049694

A:Accession: JG0164

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-279 <MOR>

A:Cross-references: GB:AF053486

Query Match 43.9%; Score 718; DB 2; Length 279;
 Best Local Similarity 42.4%; Pred. No. 9.6e-45;
 Matches 118; Conservative 46; Mismatches 112; Indels 2; Gaps 2;

QY 1 MSESFDCAKCNESLYGRKVIQTDSPGYCPVPCYDNTFANTCAEQQLIGHDSRELFPYEDRH 60
 DB 1 MSE-FKCHCEESLOGKRYVQKHCCCLCKDFCANTVCBCRPIGADSKEVHYKNRF 59

```
Qy 61 PHECFRCRCQSLADEPFTQDSSELLCNDVCYSAFSSQCSACGETVMPGSRKLEYGGQ 120
Db 60 WHNTFCQTKGSQLLATETETVAMDKNILKNKCATRVTFPKCXGCKLXIEGDHXYEYKGS 119
Qy 121 TWHEHCFCLSCGEQPLGSRFVDPDKGAHYCVPCYENKFAFRCARCSKTLTQGGVYTRDQ 180
Db 120 IWHKNCVFCTNCKDIITGKNFFPKDEGFYCVCTYDALFTKCMCKKRPITSGGVSYQDQ 179
Qy 181 WHRECLVTCGOTPLAQOQTSRDEDPYCVACFGELEFAPKCSSCKRPVIGLG-GGKYVSF 239
Db 180 WHSECFVVCSSKELSGQRTAMDDQYFCVDYKNIYAKKACGCKNITGFGKGANVAH 239
Qy 240 EDRHWHNCFSCARCSSTSLVGQGFVPDGDQVLCGCSQ 277
Db 240 EQNSWHDYCFNCKTCSVNLANKHEVFHDEQVYCPDCAR 277

RESULT 9
A55071
hydrogen peroxide-inducible protein hic-5 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: A55071
R:Shibanuma, M.; Mashimo, J.; Kuroki, T.; Nose, K.
J. Biol. Chem. 269, 26767-26774, 1994
A:Title: Characterization of the TGFbeta1-inducible hic-5 gene that encodes a putative
A:Reference number: A55071; MUID:95014536; PMID:7929412
A:Accession: A55071
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-444 <SHI>
C:Superfamily: LIM metal-binding repeat homology
F:211-261/Domain: LIM metal-binding repeat homology <LIM1>
F:270-320/Domain: LIM metal-binding repeat homology <LIM2>
F:329-379/Domain: LIM metal-binding repeat homology <LIM3>
F:388-438/Domain: LIM metal-binding repeat homology <LIM4>

Query Match 19.1%; Score 311.5; DB 2; Length 444;
Best Local Similarity 32.4%; Pred. No. 2.2e-15;
Matches 61; Conservative 23; Mismatches 89; Indels 15; Gaps 3;

Qy 101 CSACGETVMPGSRKLEYGGTWHHEHCFCLSCGEQPLGSRFVDPDKGAHYCVPCYENKFA 160
Db 211 CGSCKNKP1--AGQVVTALGRAWHPHCLSCGSTTLGSSFFERKDGAPCECYFERFSP 268
Qy 161 RCARCSKTLTQGGVYTRDQPHHRECLVTCGOTPLAQOQTSRDEDPYCVACFGELEFAPK 220
Db 269 RCGFCNQPIRHKMVTALGTHWHPHFHCCVSCGEPFGEFGEFHERGRPYCRRDFLQLFAPR 328
Qy 221 CSSCKRPVIGLGKGYVSFFEDRHHNCFSCARCSSTSLVGQGFVPDGDQVLCOG----- 274
Db 329 CQCGQGPIL----DNYISALSWHPDVCVRECLAPFGSGSFHEGRPLCENHFAQR 384
Qy 275 ---CSQAG 279
Db 385 GSLCATCG 392

RESULT 10
JC2324
LIM protein - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
C:Accession: JC2324
R:Rearden, A.
Biochem. Biophys. Res. Commun. 201, 1124-1131, 1994
A:Title: A new LIM protein containing an autoeoptope homologous to "senescent cell antigen
A:Reference number: JC2324; MUID:94296379; PMID:7517666
A:Accession: JC2324
A:Molecule type: mRNA
A:Residues: 1-314 <RELA>
A:Crogs-references: EMBL:U09284; NID:g516011; PIDN:AAA20086.1; PID:g516012
```

```
A:Experimental source: liver
C:Superfamily: LIM metal-binding repeat homology
C:Keywords: glycoprotein
F:10-62/Domain: LIM metal-binding repeat homology <LIM1>
F:71-121/Domain: LIM metal-binding repeat homology <LIM2>
F:135-184/Domain: LIM metal-binding repeat homology <LIM3>
F:149-157/Region: PINCH autoeoptope
F:193-243/Domain: LIM metal-binding repeat homology <LIM4>
F:252-303/Domain: LIM metal-binding repeat homology <LIM5>
F:87/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.7%; Score 306.5; DB 2; Length 314;
Best Local Similarity 27.4%; Pred. No. 3.9e-15;
Matches 77; Conservative 40; Mismatches 145; Indels 19; Gaps 7;

Qy 3 ESFDCAKCNESLYGRKYIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRLFLFEDRH 62
Db 34 QCFVCAQCFQFPGLFEPFGRKYCEHDFQMLFAPCCHQCGEFI--IGRVIKANNNSWH 91
Qy 63 EGCRCRCRCQSLADEPFTQDSSELLCNDVCYSAFSSQ-----CSACGETVMPGSRKLEY 117
Db 92 PECFCRDLCOEVLADIGFVNAGRHLCRPNHNRKARGLKYICQKCHAIL--DEQPLIF 149
Qy 118 GGTWHEHCFCLSCGEQPL--GSRFVDPDKGAHYCVPCYENKFAFRCARCSKTLTQGGV 175
Db 150 KNDPYHPDHFNCAKRGELTADAREL---KGELYCLPCHDKMGVPICGACRRPIEGRVYN 206
Qy 176 YRDPWHRECLVTCGOTPLAQOQTSRDEDPYCVACFGELEFAPKCSSCKRPVIGLGK 235
Db 207 AMGQKHVHFVCAKCEKPFGLHRRHYERKGLAYETHYNQLFGDVCFHCNRVI----EGD 262
Qy 236 YVSFEDRHHNCFSCARCSSTSL-VGQGFVPDGDQVLCOG 275
Db 263 VVSALNKAQVNCVFACSTCTNLTLLKNKFVDFDKMKPVCKKC 303

RESULT 11
A55933
paxillin - human
N:Alternate names: peroxisomal membrane protein 3
C:Species: Homo sapiens (man)
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 20-Apr-2000
C:Accession: A55933
R:Salgia, R.; Li, J. L.; Lo, S. H.; Brunkhorst, B.; Kansas, G. S.; Sobhany, E. S.; Sun, Y.
J. Biol. Chem. 270, 5039-5047, 1995
A:Title: Molecular cloning of human paxillin, a focal adhesion protein phosphorylated
A:Reference number: A55933; MUID:95197488; PMID:7534286
A:Molecule type: mRNA
A:Accession: A55933
A:Residues: 1-557 <SAL>
A:Crogs-references: GB:U14588; NID:g704347; PIDN:AAC50104.1; PID:g704348
C:Genetics:
A:Gene: GDB:PXN
A:Crogs-references: GDB:702105
A:Map position: 12q24-12q24
C:Superfamily: LIM metal-binding repeat homology
C:Keywords: cell adhesion; cytoskeleton; peroxisome; phosphoprotein; zinc finger
F:46-55/Region: proline-rich
F:324-374/Domain: LIM metal-binding repeat homology <LIM1>
F:383-433/Domain: LIM metal-binding repeat homology <LIM2>
F:442-492/Domain: LIM metal-binding repeat homology <LIM3>
F:501-551/Domain: LIM metal-binding repeat homology <LIM4>
F:31,118,181/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 18.6%; Score 304.5; DB 2; Length 557;
Best Local Similarity 27.5%; Pred. No. 8.4e-15;
Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3;

Qy 40 CAEQQLIGHDSRLFLFEDRHFGCFRCRCRCQSLADEPFTQDSSELLCNDVCYSAFSS 99
Db 324 CGACKKP1A--GQVVTAMGTWHPHFVCTHCQEEIGSRNFFERDGGPYCEKDYHNLFSP 381
Qy 100 QCSACGETVMPGSRKLEYGGQTWHEHCFCLSCGEQPLGSRFVDPDKGAHYCVPCYENKFA 159
```


Db 382 RCYYCNGPIL--DKVVTALDRTWHPHEFFCAQCGAFFGPEGFHEKDGKAYCRKDYFDMEA 439
QY 160 PRCARCSKTLTQGGVYTRDQPWRECLVCTGCCTPLAGQOFTSRDEDPPYCVACFGELFAP 219
Db 440 PKCGGCARAILENYISALNTLWHPFVCEFCVCRECFPFVNGSFFFDHGDQPYCEVHYHERRGS 499
QY 220 KCSCKRPVIGLGGGKVVSPEDRHHNHCFCSCARCTSLVGGQGVFPDQGVLCQGC 275
Db 500 LCSCGCKPIT---GRCITAMGKFFHPEHFVCAFLKQLNKGTFFKEQNDKPYCQNC 551

RESULT 12
B55933
paxillin - chicken.
C:Species: Gallus gallus (Chicken)
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 21-Jul-2000
C:Accession: B55933
R:Salgia, R.; Li, J.L.; Lo, S.H.; Brunkhorst, B.; Kansas, G.S.; Sobhany, E.S.; Sun, Y.; J. Biol. Chem. 270, 5039-5047, 1995
A:Title: Molecular cloning of human paxillin, a focal adhesion protein phosphorylated by A:Reference number: A55933; MUID:95197488; PMID:7534286
A:Accession: B55933
A:Molecule type: mRNA
A:Residues: 1-559 <SAL>
A:Cross-references: GB:U14589; NID:g704349; PIDN:AAC59665.1; PID:g704350
C:Superfamily: LIM metal-binding repeat homology
C:Keywords: cell adhesion; cytoskeleton; phosphoprotein; zinc finger
F:46-55/Region: proline-rich
F:326-376/Domain: LIM metal-binding repeat homology <LIM1>
F:385-435/Domain: LIM metal-binding repeat homology <LIM2>
F:444-494/Domain: LIM metal-binding repeat homology <LIM3>
F:503-553/Domain: LIM metal-binding repeat homology <LIM4>
F:31,118/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 18.6%; Score 303.5; DB 2; Length 559;
Best Local Similarity 27.9%; Pred. No. 9.9e-15;
Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3;

QY 40 CABCQQLIGHDSRELFEVDHRHFHGGCFRCRCRCQSLADEPFTCDSELLCNDCCYSAFSS 99
Db 326 CGACKRPIA--GOVVTAMGKTWHPHFVCTHCQEEIGSRNFFERDGPCEKDYHNLFSP 383
QY 100 QCSACGETWPMGRKLEYGGQTWHEHCFHLCGCEQPLGSRSFVPDKGAHYCVPCYENKFA 159
Db 384 RCYYCNGPIL--DKVVTALDRTWHPHEFFCAQCGVFFGPEGFHEKDGKAYCRKDYFDMEA 441
QY 160 PRCARCSKTLTQGGVYTRDQPWRECLVCTGCCTPLAGQOFTSRDEDPPYCVACFGELFAP 219
Db 442 PKCGGCARAILENYISALNTLWHPFVCEFCVCRECFPFVNGSFFFDHGDQPYCEVHYHERRGS 501
QY 220 KCSCKRPVIGLGGGKVVSPEDRHHNHCFCSCARCTSLVGGQGVFPDQGVLCQGC 275
Db 502 LCSCGCKPIT---GRCITAMGKFFHPEHFVCAFLKQLNKGTFFKEQNDKPYCQNC 553

RESULT 13
E88469
protein C28H8 6 [imported] - Caenorhabditis elegans.
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: E88469
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: published genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele
A:Note: see bibliated errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: E88469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-256 <STO>
A:Cross-references: GB:chr_III; PIDN:AAA62294.1; PID:g669012; GSPDB:GNO00021; CESP:C28H8

RESULT 13
E88469
protein C28H8.6 [imported] - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: E88469
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biological processes
A:Reference number: A75000; PMID:99069633; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2104, 1999.
A:Accession: E88469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-256 <STO>
A:Cross-references: GB:chr_III; PIDN:AAA62294.1; PID:g669012; GSPDB:GNO00021; CESP:C28H8.6

```
Dbb      382 RCYYCNGPIL--DKVVTALDRTPWHPEFTCAQCGAFFGPGFHEKGKAYCRKDYDFDMEA   439
          :| |:: :: |||| ||:|| | | | | | | | | | | | | | | | | | | | | | |
Qy       160 PRCARCSKTQTGGVVTRDPWPWRECLVCTGCOTPLAGQQOFTSRDESDPYCVACFGELFPAP    219
          |:|:: :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db        440 PKGGCCARALLENVISALNTLWNPECFVCRECFTPFVNGSFFEHGDGPCEVHYHERRGS     499
          |||:: :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy       220 KSCSCKRPVGLGGGVKVSFPEDRHHNHCFCARCSTSLVGQGVPFDGQVLQCQC           275
          |||::::|:: :| | | | | | | | | | | | | | | | | | | | | | | | | |
Db        500 LCSGCCOKPII---GRCIITAMAKFKPEHFVFCAFCLKQLNKGTTFKEONDKPYCONC    551
          |||::::|:: :| | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 12
B55933
paxillin - chicken.
C;Species: Gallus gallus (chicken)
C;Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 21-Jul-2000
C;Accession: B55933
E;Salgia, R.; Li, J.L.; Lo, S.H.: Brunkhorst, B.; Kansas, G.S.; Sobhani, E.S.; Sun, Y.,
J. Biol. Chem. 270, 5039-5047, 1995
A;Title: Molecular cloning of human paxillin, a focal adhesion protein phosphorylated by
A;Reference number: A55933; PMID:95197488; PMID:7534286
A;Accession: B55933

[illegible]

RESULT 13
E88469
protein C28H8.6 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: E88469
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
A:Reference number: A75000; PMID:9851916
A:Note: See vbsites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: Published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2104, 1999.
A:Accession: E88469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-256 <STO>
A:Cross-references: GB:chr_III; PIDN:AAA622294.1; PID:g669012; GSPDB:GNO00021; CESP:C28H8.6

Db	212	TEERVIALGKHWHVEHVCVCKBPKFLGHHHRTKGLPYCEQHFKHLFGNLCTKCGDPC	271
QY	229	VLGGGKYVSFEDRHHHNCFSARCSTSL	258
Db	272	C---GEVFOALOKTWCVKCFSCFCDDKL	297

RESULT 15

N5501A 1J
 A55050
 enigma - human
 C: Species: Homo sapiens (man)
 C: Date: 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
 C: Accession: A55050
 J: Wu, R.Y.; Gill, G.N.
 J: Biol. Chem. 269, 25085-25090, 1994
 A: Title: LIM domain recognition of a tyrosine-containing tight turn.
 A: Reference number: A55050; MUID:95014287; PMID:7929196
 A: Accession: A55050
 A: Status: preliminary
 A: Molecule type: mRNA
 A: Residues: 1-455 <WUA>
 A: Cross-references: GB:L35240; NID:g561636; PIDN:AAC37565.1; PID:g561637
 C: Superfamily: LIM metal-binding repeat homology; GLGF domain homology
 F: 81-330/Domain: LIM metal-binding repeat homology <GLGF>
 F: 280-330/Domain: LIM metal-binding repeat homology <LIM1>
 F: 339-389/Domain: LIM metal-binding repeat homology <LIM2>
 F: 398-450/Domain: LIM metal-binding repeat homology <LIM3>

	Query Match	17.9%	Score 293;	DB 2;	Length 455;
	Best Local Similarity	32.4%;	Pred. No. 4.8e-14;		
	Matches	56;	Conservative 27;	Mismatches 86;	Indels 4; Gaps 2;
QY	101	CSACGETVMPGSRKLEYGQGTWHEHCFCLSCGCEQPLGSRSFVDPDKGNAHYCVPCYENK	FAP	160	
Db	280	CHQCHKVR--GRYLVALGHAYHPEEFVCSQCKVLEEGGFBEKGAIFCCPYCDYRYAP		337	
QY	161	RCARCSKTLTQGGVTYRDOPWHRECLVCTGCQPTLAGQQFTSDSDPYCVACFGEI	FAPK	220	
Db	338	SCAKCKKKITGETIMHALKWTWHVHCFTCAACKTPIRNRAFYMEEGPYPCERDEYKMG	ETK	397	
QY	221	CSSCKRPTVGLGGKXVSPEDRHHHNCFSARCSTSLVGQGVPPGDQVLQC		273	
Db	398	CHGCDPKID--AGDREALGFWSHDTCFVCAICQINLEKTYFSYKKDRPLCK		448	

Search completed: August 22, 2003, 20:00:42
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:44:56 ; Search time 23 Seconds
(without alignments)
572.499 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSESFDCARKNESLYGRKYI.....QGFPDQVLCQCSQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt.41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1578.5	96.5	289	1 SLI2_MOUSE	Q9R059 mus musculus
2	1569	96.0	280	1 SLI2_HUMAN	Q13643 homo sapien
3	998	61.0	279	1 SLI3_RAT	Q35115 rattus norv
4	991	60.6	279	1 SLI3_MOUSE	O70433 mus musculus
5	968	59.2	279	1 SLI3_HUMAN	Q14192 homo sapien
6	762.5	46.6	280	1 SLI1_HUMAN	Q13642 mus musculus
7	760.5	46.5	280	1 SLI1_MOUSE	P57447 mus musculus
8	313	19.1	386	1 LPAX_HUMAN	O60711 homo sapien
9	309.5	18.9	325	1 PINC_HUMAN	P48059 mus sapien
10	305	18.7	495	1 MLP2_DROME	Q24400 drosophila
11	304.5	18.6	591	1 PAXI_HUMAN	P49023 homo sapien
12	303.5	18.6	559	1 PAXI_CHICK	P49024 gallus gall
13	300.5	18.4	256	1 YP96_CAEEL	O09476 caenorhabdi
14	297.5	18.2	348	1 UN97_CAEEL	P50464 caenorhabdi
15	265	16.2	615	1 LM06_HUMAN	Q43900 homo sapien
16	257.5	15.7	572	1 ZYX_HUMAN	Q15942 mus sapien
17	254	15.5	542	1 ZYX_CHICK	Q04584 gallus gall
18	244.5	15.0	564	1 ZYX_MOUSE	O62523 mus musculus
19	240	14.7	476	1 TRI6_HUMAN	Q15654 homo sapien
20	237	14.5	329	1 PIN2_CAEEL	P19157 caenorhabdi
21	226	13.8	423	1 TES_MOUSE	P47226 mus musculus
22	220	13.5	421	1 TES_HUMAN	Q90418 homo sapien
23	200	12.2	194	1 CSR3_HUMAN	P50461 homo sapien
24	196.5	12.0	192	1 CSR2_HUMAN	Q16527 homo sapien
25	196.5	12.0	365	1 LMCI_HUMAN	Q9NZU5 homo sapien
26	195.5	12.0	192	1 CSR2_RAT	Q62908 rattus norv
27	195	11.9	193	1 CSR2_CHICK	P50460 gallus gall
28	195	11.9	193	1 CSR2_COTJA	Q05158 coturnix co
29	195	11.9	194	1 CSR3_MOUSE	P50462 mus musculus
30	194	11.9	194	1 CSR3_RAT	P50463 rattus norv
31	194	11.9	365	1 LMCI_MOUSE	Q8VEE1 mus musculus
32	182	11.1	192	1 CYSR_RAT	P47875 rattus norv
33	182	11.1	898	1 Z071_XENLA	P18751 xenopus lae

RESULT 1

ID	SLI2_MOUSE	STANDARD;	PRT;	289 AA.
AC	Q9R059; Q9JLP5; Q9WUH3;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Skeletal muscle LIM-protein 2 (SLIM 2) (Four and a half LIM domains			
DE	protein 3) (FHL-3).			
GN	FHL3.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RP	MEDLINE=99160848; PubMed=10049693;			
RA	Morgan M.J., Madgwick A.J.A.;			
RT	"The LIM proteins FHL1 and FHL3 are expressed differently in skeletal			
RT	muscle.";			
RL	Biochem. Biophys. Res. Commun. 255:245-250(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RP	MEDLINE=20368180; PubMed=10906474;			
RA	Chu P.-H., Ruiz-Lozano P., Zhou Q., Cai C., Chen J.;			
RT	"Expression patterns of FHL/SLIM family members suggest important			
RT	functional roles in skeletal muscle and cardiovascular system.";			
RL	Mech. Dev. 95:259-265(2000).			
RN	[3]			
RP	SEQUENCE OF 11-228 FROM N.A.			
RA	Li H.Y., Lee S.M.Y., Tsui S.K.W., Chan K.K., Kotaka M., Chim S.S.C.,			
RA	Lee C.Y., Fung K.P., Waye M.M.Y.;			
RT	"The cloning, sequencing and characterization of a mouse FHL3, which			
RT	contains four and a half LIM domains.";			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
CC	-1- TISSUE SPECIFICITY: In the adult, expressed at high levels in			
CC	skeletal muscle and, to a lesser extent, in heart, lung, skin and			
CC	kidney. During embryonic development, expressed ubiquitously at			
CC	low levels.			
CC	-1- SIMILARITY: Contains 4 LIM zinc-binding domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed, usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF114382; AAD53231.1; -			
DR	EMBL; AF134772; AAD32623.1; -			
DR	EMBL; AF149826; AAF73159.1; -			
DR	HSSP; P04006; 1IML.			
DR	MGD; MGI:1341092; Fhl3.			
DR	InterPro; IPR001781; LIM.			
DR	Pfam; PF00412; LIM; 4.			

34	179.5	11.0	3075	1	LMAL_HUMAN	P25391 homo sapien
35	178	10.9	192	1	CYSR_HUMAN	P21291 homo sapien
36	177	10.8	191	1	CYSR_CHICK	P32965 gallus gall
37	174.5	10.7	642	1	LIK2_CHICK	P53666 gallus gall
38	170.5	10.4	606	1	ZRG6_XENLA	P18733 xenopus lae
39	169.5	10.4	1017	1	LRG1_YEAST	P35688 saccharomyc
40	167	10.2	208	1	CRP2_HUMAN	P52943 homo sapien
41	166.5	10.2	647	1	LIK1_RAT	P53669 rattus norv
42	166	10.2	208	1	CRP2_RAT	P36201 rattus norv
43	163.5	10.0	398	1	LHX3_BRARE	Q30421 brachydanio
44	163	10.0	367	1	LHX4_HUMAN	Q96992 homo sapien
45	163	10.0	367	1	LHX4_MOUSE	P53776 mus musculu

ALIGNMENTS

DR ProDom; PD000094; LIM; 4.
 DR SMART; SM00132; LIM; 4.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 3.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 4.
 KW Repeat; LIM domain; Metal-binding; Zinc; Zinc-finger.
 FT ZN_FING 7 31 GATA-LIKE (POTENTIAL).
 FT DOMAIN 40 92 LIM 1.
 FT DOMAIN 101 153 LIM 2.
 FT DOMAIN 162 212 LIM 3.
 FT DOMAIN 221 275 LIM 4.
 FT CONFLICT 77 77 D -> G (IN REF. 1).
 FT CONFLICT 91 91 E -> D (IN REF. 1).
 FT CONFLICT 95 95 T -> S (IN REF. 1).
 FT CONFLICT 137 137 A -> G (IN REF. 1 AND 3).
 FT CONFLICT 157 159 KFA -> NLT (IN REF. 3).
 FT CONFLICT 192 192 K -> Q (IN REF. 1).
 FT CONFLICT 199 199 H -> Q (IN REF. 1 AND 3).
 FT CONFLICT 205 205 D -> E (IN REF. 1).
 FT CONFLICT 225 225 N -> K (IN REF. 1 AND 3).
 FT CONFLICT 229 238 TGGSGGAEGA -> V (IN REF. 1).
 FT CONFLICT 256 256 S -> N (IN REF. 1).
 SQ SEQUENCE 289 AA; 31817 MW; 176DD857FCD1B9D9 CRC64;

Query Match 96.5%; Score 1578.5; DB 1; Length 289;
 Best Local Similarity 93.4%; Pred. No. 1.8e-122;
 Matches 270; Conservative 6; Mismatches 4; Indels 9; Gaps 1;

QY 1 MSESFDCAKNEISLYGRKVIQTDSPGVCPCVDNTFANTCAECQQLIGHDSRELFEYDRH 60
 |||||
 DB 1 MSEAEDCAKNEISLYGRKVIQTDSPGVCPCVDNTFANTCAECQQLIGHDSRELFEYDRH 60
 |||||

QY 61 FHEGFCRCRCORSIADPEFTQDSSELCNDCYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
 |||||
 DB 61 FHEGFCRCRCORSIADPEFTQDSSELCNDCYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
 |||||

QY 121 TWHEHCFLCSCGEQPLGSRFVDPKGAHYCPVCYENKFAPRCARCKTLTGGVTVYRDQP 180
 |||||
 DB 121 TWHEHCFLCSCGEQPLGSRFVDPKGAHYCPVCYENKFAPRCARCKTLTGGVTVYRDQP 180
 |||||

QY 181 WHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPITGGGGAGAGL 240
 |||||
 DB 181 WHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPITGGGGAGAGL 240
 |||||

QY 232 GGGKYVSFEDRHHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAP 280
 |||||
 DB 241 GGGKYVSFEDRHHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAP 289
 |||||

RESULT 2

SLI2_HUMAN
 ID SLI2_HUMAN STANDARD; PRT; 280 AA.
 AC Q13643.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Skeletal muscle LIM-protein 2 (SLIM 2) (Four and a half LIM domains
 DE protein 3) (FHL-3).
 GN FHL3 OR SLIM2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=99160848; PubMed=10049693;
 RA Morgan M.J., Madgwick A.J.A.;
 RT "The LIM proteins FHL1 and FHL3 are expressed differently in skeletal
 RT muscle.";
 RL Biochem. Biophys. Res. Commun. 255:245-250(1999).
 RN [2]
 RP PRELIMINARY SEQUENCE OF 16-280 FROM N.A.
 RC TISSUE=Skeletal muscle;

RX MEDLINE=96354835; PubMed=8753811;
 RA Morgan M.J., Madgwick A.J.A.;
 RT "Slim defines a novel family of LIM-proteins expressed in skeletal
 RT muscle.";
 RL Biochem. Biophys. Res. Commun. 225:632-638(1996).
 CC -!- TISSUE SPECIFICITY: EXPRESSED ONLY IN SKELETAL MUSCLE.
 CC -!- SIMILARITY: Contains 4 LIM zinc-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U60116; AAC04466.2; -.
 DR PIR; T09504; T09504.
 DR HSSP; P32965; ICTL.
 DR Gene; HGNC:3704; FHL3.
 DR MIM; 602790; -.
 DR GO; GO:0007517; P:muscle development; TAS.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 4.
 DR ProDom; PD000094; LIM; 4.
 DR SMART; SM00132; LIM; 4.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 4.
 KW Repeat; LIM domain; Metal-binding; Zinc; Zinc-finger.
 FT ZN_FING 7 31 GATA-LIKE (POTENTIAL).
 FT DOMAIN 40 92 LIM 1.
 FT DOMAIN 101 153 LIM 2.
 FT DOMAIN 162 212 LIM 3.
 FT DOMAIN 221 275 LIM 4.
 SQ SEQUENCE 280 AA; 31291 MW; 21AFF0950045A3DC CRC64;

Query Match 96.0%; Score 1569; DB 1; Length 280;
 Best Local Similarity 95.7%; Pred. No. 1e-121;
 Matches 268; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSESFDCAKNEISLYGRKVIQTDSPGVCPCVDNTFANTCAECQQLIGHDSRELFEYDRH 60
 |||||
 DB 1 MSESFDCAKNEISLYGRKVIQTDSPGVCPCVDNTFANTCAECQQLIGHDSRELFEYDRH 60
 |||||

QY 61 FHEGFCRCRCORSIADPEFTQDSSELCNDCYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
 |||||
 DB 61 FHEGFCRCRCORSIADPEFTQDSSELCNDCYCSAFSSQCSACGETVMPGSRKLEYGGQ 120
 |||||

QY 121 TWHEHCFLCSCGEQPLGSRFVDPKGAHYCPVCYENKFAPRCARCKTLTGGVTVYRDQP 180
 |||||
 DB 121 TWHEHCFLCSCGEQPLGSRFVDPKGAHYCPVCYENKFAPRCARCKTLTGGVTVYRDQP 180
 |||||

QY 181 WHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPITGGGGAGAGL 240
 |||||
 DB 181 WHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGELFAPKCSCKRPITGGGGAGAGL 240
 |||||

QY 241 DRHHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAP 280
 |||||
 DB 241 DRHHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAP 280
 |||||

RESULT 3
 SLI3_RAT
 ID SLI3_RAT STANDARD; PRT; 279 AA.
 AC O35115;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL)
 DE (Four and a half LIM domains protein 2) (FHL-2).
 GN FHL2 OR SLIM3 OR DRAL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RA Tanahashi H.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 4 LIM zinc-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB008571; BAA23357.1; -
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 4.
 DR ProDom; PD000094; LIM; 4.
 DR SMART; SM00132; LIM; 4.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 4.
 KW Repeat; LIM domain; Metal-binding; zinc; Zinc-finger.
 FT ZN_FING 7 31 GATA-LIKE (POTENTIAL).
 FT DOMAIN 40 92 LIM 1.
 FT DOMAIN 101 153 LIM 2.
 FT DOMAIN 162 212 LIM 3.
 FT DOMAIN 221 275 LIM 4.
 SQ SEQUENCE 279 AA; 32086 MW; 9A9D8E5935034173 CRC64;
 Query Match 61.0%; Score 998; DB 1; Length 279;
 Best Local Similarity 55.6%; Pred. No. 5.7e-75;
 Matches 154; Conservative 55; Mismatches 68; Indels 0; Gaps 0;
 QY 1 MSSEFDCAKNESLYGRKYIQTDSGPYCPYDNTFANTCAEQQLIGHDSRELFLYEDRH 60
 DB 1 MTERFDCHNCNESLYGKYLKEENPHCVACFEELYANTCECGTPIGCDCKDLSYKDRH 60
 QY 61 FHGCFRCRCQSLADEPTCODSELLCNDYCSAFSSQCSACGETVMPGSKRYEGGQ 120
 DB 61 FHGCFRCRCQSLADEPTCODSELLCNDYCSAFSSQCSACGETVMPGSKRYEGGQ 120
 QY 121 TWHEHCFVLCGCGQPLAGQFTSRDEPDYCVACFGELFAPKCSCKRPVGLGGKYSVFE 180
 DB 121 SWHETCFRCRCQSLADEPTCODSELLCNDYCSAFSSQCSACGETVMPGSKRYEGGQ 180
 QY 181 WHRECLVCTGCTPLAGQFTSRDEPDYCVACFGELFAPKCSCKRPVGLGGKYSVFE 240
 DB 181 WHRECFVCTACKQLSGQRTARDEFPYCLTFCFDLYAKKACGCTNPISGLGCTKYSFE 240
 QY 241 DRWHHNCFCARCSTSLVGGQFVPGDQVLCGCGSQ 277
 DB 241 ERQWHDNCFNCCKCSLSVGRGFLTERDDILCPDCGK 277
 RESULT 4
 SLI3_MOUSE STANDARD; PRT; 279 AA.
 ID 070433; P97448;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL)
 DE (Four and a half LIM domains protein 2) (FHL-2).
 GN FHL2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Brain;
 RA Chan K.K., Tsui S.K.W., Lee C.Y., Fung K.P., Waye M.M.Y.;
 RT "The cloning, sequencing and characterization of a mouse FHL2, which
 RT contains four and a half LIM domains.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE-99160848; PubMed-10049693;
 RA Morgan M.J., Madgwick A.J.A.;
 RT "The LIM proteins FHL1 and FHL3 are expressed differently in skeletal
 RT muscle.";
 RL Biochem. Biophys. Res. Commun. 255:245-250(1999).
 RN [3]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE-20368180; PubMed-10906474;
 RA Chu P.-H., Ruiz-Lozano P., Zhou Q., Cai C., Chen J.;
 RT "Expression patterns of FHL/SLIM family members suggest important
 RT functional roles in skeletal muscle and cardiovascular system.";
 RL Mech. Dev. 95:259-265(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RA Starzinski-Powitz A., Martin B., Eckardt F.;
 RT "Isolation of the mouse homolog mDRAL from skeletal muscle derived
 RT myoblasts.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- TISSUE SPECIFICITY: Highly expressed in heart but also detectable
 CC in brain and skeletal muscle.
 CC -1- SIMILARITY: Contains 4 LIM zinc-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF055889; AAC12770.1; -
 DR EMBL; U77040; AAB19211.2; -
 DR EMBL; AF114381; AAD53230.1; -
 DR EMBL; AF153340; AAD34170.1; -
 DR MGD; MGI:1338762; Fhl2.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0005515; P:protein binding activity; IPI.
 DR GO; GO:0006357; P:regulation of transcription from Pol II pro. . . ; IDA.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 4.
 DR ProDom; PD000094; LIM; 4.
 DR SMART; SM00132; LIM; 4.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 4.
 KW Repeat; LIM domain; Metal-binding; zinc; Zinc-finger.
 FT ZN_FING 7 31 GATA-LIKE (POTENTIAL).
 FT DOMAIN 40 92 LIM 1.
 FT DOMAIN 101 153 LIM 2.
 FT DOMAIN 162 212 LIM 3.
 FT DOMAIN 221 275 LIM 4.
 SQ SEQUENCE 279 AA; 32072 MW; 6D8CBC4B424BFF2 CRC64;
 Query Match 60.6%; Score 991; DB 1; Length 279;
 Best Local Similarity 54.9%; Pred. No. 2.1e-74;
 Matches 152; Conservative 57; Mismatches 68; Indels 0; Gaps 0;
 QY 1 MSSEFDCAKNESLYGRKYIQTDSGPYCPYDNTFANTCAEQQLIGHDSRELFLYEDRH 60
 DB 1 MTERFDCHNCNESLYGKYLKEENPHCVACFEELYANTCECGTPIGCDCKDLSYKDRH 60
 QY 61 FHGCFRCRCQSLADEPTCODSELLCNDYCSAFSSQCSACGETVMPGSKRYEGGQ 120
 DB 61 FHGCFRCRCQSLADEPTCODSELLCNDYCSAFSSQCSACGETVMPGSKRYEGGQ 120
 QY 121 TWHEHCFVLCGCGQPLAGQFTSRDEPDYCVACFGELFAPKCSCKRPVGLGGKYSVFE 180
 DB 121 SWHETCFRCRCQSLADEPTCODSELLCNDYCSAFSSQCSACGETVMPGSKRYEGGQ 180
 QY 181 WHRECLVCTGCTPLAGQFTSRDEPDYCVACFGELFAPKCSCKRPVGLGGKYSVFE 240
 DB 181 WHRECFVCTACKQLSGQRTARDEFPYCLTFCFDLYAKKACGCTNPISGLGCTKYSFE 240
 QY 241 DRWHHNCFCARCSTSLVGGQFVPGDQVLCGCGSQ 277
 DB 241 ERQWHDNCFNCCKCSLSVGRGFLTERDDILCPDCGK 277

QY 241 DRWHNHCSCARCTSLVGGQFVDPDQVLCQGCQ 277
 Db 241 ERQWHDNCFNCKKCSLSLVGRGFLTERDDILCPDCGK 277

RESULT 6

SLIM_HUMAN STANDARD; PRT; 280 AA.
 AC Q13642; Q13645;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Skeletal muscle LIM-protein 1 (SLIM 1) (SLIM) (Four and a half LIM
 domains protein 1) (FHL-1).
 GN FHL1 OR SLIM1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=96354835; PubMed=8753811;
 RA Morgan M.J., Madgwick A.J.A.;
 RT "Slim defines a novel family of LIM-proteins expressed in skeletal
 muscle.";
 RL Biochem. Biophys. Res. Commun. 225:632-638(1996).
 RN [2]
 RP REVISIONS.
 RA Morgan M.J.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=98382585; PubMed=9714789;
 RA Lee S.M.Y., Tsui S.K.W., Chan K.K., Garcia-Barcelo M., Waye M.M.Y.,
 RA Fung K.P., Liew C.C., Lee C.Y.;
 RT "Chromosomal mapping, tissue distribution and cDNA sequence of
 four-and-a-half LIM domain protein 1 (FHL1).";
 RL Gene 216:163-170(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Chapman J.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usslin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP SEQUENCE OF 106-255 FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE=95352104; PubMed=7626119;
 RA Morgan M.J., Madgwick A.J.A., Charlestone B., Pell J.M., Loughna P.T.;

RT "The developmental regulation of a novel muscle LIM-protein.";
 RL Biochem. Biophys. Res. Commun. 212:840-846(1995).
 CC -1- FUNCTION: MAY HAVE AN INVOLVEMENT IN MUSCLE DEVELOPMENT OR
 CC HYPERTROPHY.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SKELETAL MUSCLE AND TO A
 CC LESSER EXTENT IN PLACENTA, OVARY, PROSTATE, TESTIS, SMALL
 CC INTESTINE, COLON AND SPLEEN. NOT EXPRESSED IN BRAIN, LUNG, LIVER,
 CC KIDNEY, PANCREAS, THYMUS AND PERIPHERAL BLOOD LEUKOCYTES.
 CC -1- DEVELOPMENTAL STAGE: ELEVATED LEVELS DURING POSTNATAL MUSCLE
 CC GROWTH.
 CC -1- SIMILARITY: Contains 4 LIM zinc-binding domains.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U60115; AAC52021.1; -;
 DR EMBL; U29538; AAC35421.1; -;
 DR EMBL; AL078638; CAC18881.1; -;
 DR EMBL; BC010998; AAH10998.1; -;
 DR EMBL; U60118; AAC50795.1; -;
 DR HSSP; P32965; 1CTL.
 DR Genew; HGNC:3702; FHL1.
 DR MIM; 300163; -;
 DR GO; GO:0016049; P:cell growth; NAS.
 DR GO; GO:0007517; P:muscle development; NAS.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 4.
 DR ProDom; PD000094; LIM; 4.
 DR SMART; SM00132; LIM; 4.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
 DR PROSITE; PS00033; LIM_DOMAIN_2; 4.
 DR Repeat; LIM domain; Metal-binding; Zinc; Developmental protein;
 KW Differentiation; Zinc-finger.
 FT ZN.FING 7 31 GATA-LIKE (POTENTIAL).
 FT DOMAIN 40 92 LIM 1.
 FT DOMAIN 101 153 LIM 2.
 FT DOMAIN 162 212 LIM 3.
 FT DOMAIN 221 276 LIM 4.
 FT CONFLICT 73 73 H -> Q (IN REF. 1).
 FT CONFLICT 98 98 S -> F (IN REF. 1).
 FT CONFLICT 158 158 F -> L (IN REF. 1 AND 6).
 FT SEQUENCE 280 AA; 31895 MW; 2FC873D70E62834D CRC64;
 Query Match 46.6%; Score 762.5; DB 1; Length 280;
 Best Local Similarity 44.2%; Pred. No. 1.le-55;
 Matches 123; Conservative 52; Mismatches 102; Indels 1; Gaps 1;
 QY 1 MSFSPFCACNCSLYGRKYIQTDSGPYCPYDNTFANTCAEQQLIGHDSRELFYEDRH 60
 Db 1 MAEFPCYCHYCRDPLQKQYVQKDGHCCLCFKFCANTCVCCKPKTGADSKVEHYKNRF 60
 QY 61 FHEGFCRCRCORSLADEPPTCDSELLCNDYCSAFSSOCACGETVMPGSRKLEYGGQ 120
 Db 61 WHDTCFCAKCLHPLANETFAKNDKILCNKCTREDSPKCKGCFRAIVAGDQNVYKGT 120
 QY 121 TWHEHCFCLSCGBOPLGSRFSFVDPKGAHYCPVCYENKFAFPCARCSKTLFGGVYRDQP 180
 Db 121 VHKDCFTCSNCKQVIGTGSPFKGDFYCVTCHETKFAKHKVCNKAITSGGITVYDQP 180
 QY 181 WHRECLVCTCQTPLAGQQFTSRDEDDYCVACGFGELFAPKCSCKRPIVGLG-GGKYVSF 239
 Db 181 WHADCFVCTVCSKKLAGQRFATVEDQYVCYDCKNFVAKKCAKCKNPITGFGKSSVYAY 240
 QY 240 EDRHWHNCFSCARCTSLVGGQFVDPDQVLCQGCQ 277
 Db 241 EGSHWDHCFCHCKKCSVNLANKRFFVHQEQVCPDCAK 278

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: PLAYS A ROLE IN CELL DIFFERENTIATION LATE IN MYOGENESIS.
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSION IS RESTRICTED TO THE
CC SOMATIC, VICERAL, AND PHARYNGEAL MUSCLES. WITHIN THE SOMATIC
CC MUSCULATURE, MLP48B IS LOCALIZED AT THE ENDS OF MUSCLES FIBERS AT
CC THE POINT OF ATTACHMENT TO THE EPIDERMIS. THERE IS NO EXPRESSION
CC IN CARDIAC MESODERM OR IN FAT BODY.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS BIPHASIC, PEAKING LATE IN
CC EMBRYOGENESIS (16-24 H EMBRYOS) AND DURING THE LARVAL TO PUPAL
CC TRANSITION, WHEN THE MUSCULATURE IS DIFFERENTIATING. FOUND IN
CC DEVELOPING MUSCLES OF THE VISCERAL AND SOMATIC MESODERM SUBSEQUENT
CC TO THE FORMATION OF THE MUSCLE PRECURSOR CELLS. DECREASED LEVELS
CC ARE STILL DETECTABLE IN ADULTS.
CC -1- SIMILARITY: TO THE VERTEBRATE CYSTEINE-RICH PROTEINS.
CC -1- SIMILARITY: Contains 5 LIM zinc-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X91245; CAA62627.1; -;
DR EMBL; AF090832; AAC61591.1; -;
DR EMBL; AE003672; AAF54063.1; -;
DR HSSP; P32965; 1CTL.
DR FlyBase; FBgn0014863; MLP84B.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 5.
DR ProDom; PD000094; LIM; 5.
DR SMART; SM00132; LIM; 5.
DR PROSITE; PS00478; LIM_DOMAIN_1; 5.
DR PROSITE; PS00023; LIM_DOMAIN_2; 5.
DR Nuclear protein; Repeat: LIM domain; Metal-binding; Zinc; Myogenesis;
KW Developmental protein; Differentiation.
FT DOMAIN 12 63
FT DOMAIN 65 80
FT DOMAIN 66 71
FT DOMAIN 120 172
FT DOMAIN 175 180
FT DOMAIN 178 189
FT DOMAIN 222 274
FT DOMAIN 276 291
FT DOMAIN 325 377
FT DOMAIN 379 390
FT DOMAIN 421 473
FT DOMAIN 475 490
SQ SEQUENCE 495 AA; 53525 MW; 2E559B9178E54C0E CRC64;

Query Match 18.7%; Score 305; DB 1; Length 495;
Best Local Similarity 21.6%; Pred. No. 51e-18;
Matches 97; Conservative 40; Mismatches 121; Indels 192; Gaps 20;
QY 5 FDCAKNESLYGRKYIQTDSGGPYCVPCYDNTF-----ANT----- 39
DB 37 FKCGMCKSLDSTNCTEHERELYCKTCHGRKFGPKGYGFGTGLSDMNGSQFLRENGD 96
QY 40 -----CAEQQLIGHDSRELFFYEDRHFHEGCFRCRCQBSLA 76
DB 97 VPSVRNGARLEPRAIARAPEGCGPCRGGVY-YAAEQMLARGSWHKCFCKGCTCKRGL- 154
QY 77 DEFTTCQ--DSELLCNDVCYSAFSSQ----- 100
DB 155 DSILCEAPDKNIYCKGYAKKFGPKGYGVGGGALQDCYAHDDGAPAIRAIDVDKI 214
QY 101 -----CSAGETVMPGSRKLEYGGTWHFHEFLSCGCGPGLGS--RSVPDKGAIYCV 151
DB 215 QARPGGCGPCRGGVYAAEQKLS-KGREWHKCFNCCKDKHTLDSINASDPDRDV-YCR 272
QY 152 PCYENKFAPR-----SSCK-----RPVGV--LGGKYVSF-----EDRHHNCF 249
DB 273 TCYCKWPGHYGACGSGFLQDTGLTDEOISANRPYNPDTTSIKARDGEGCPRC- 328
QY 170 TOGGVTY-----RDQPHRECLVCTGCTPL-AGQOFTSRDEDPYCVACFGLFAPKC 221
DB 329 --GGAVFAAQQLSKGKWHKCYNCADCHRLDSVLACDPDGDHCRACYGKLFPGKG 386
QY 222 -----SSCK-----RPVGV--LGGKYVSF-----EDRHHNCF 249
DB 387 FGYGHAPTLVSTSGESTIQPDGRPLAGPKTSGGCPRGFAVFAAEQMIKTRTHKRCF 446
QY 250 SCARCSTSL-----VGQGFVDPGDVLCQGC 275
DB 447 YCSDCKSLDSTNLNDG--PDGD-IYCRAC 473
RESULT 11
PAXL_HUMAN STANDARD; PRT; 591 AA.
AC P49023; O14970; O14971; 060360;
DT 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Paxillin.
GN PAX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX MEDLINE=95197488; PubMed=7534286;
RA Salgia R., Li J.-L., Lo S.H., Brunkhorst B., Kansas G.S.,
RA Sobhany E.S., Sun Y., Pisick E., Hallek M., Ernst T., Tantravahi R.,
RA Chen L.B., Griffin J.D.;
RT "Molecular cloning of human paxillin, a focal adhesion protein
RT phosphorylated by P210BCR/ABL."
RL J. Biol. Chem. 270:5039-5047(1995).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RA Yamagata K., Oda N., Furuta H., Vaxillaire M., Southam L., Boriraj V.,
RA Chen X., Oda Y., Takeda J., Yamada S., Nishigori H., Lebeau M.M.,
RA Lathrop M., Cox R.D., Bell G.I.;
RT "transcription map of the 5cM region surrounding the hepatocyte
RT nuclear factor-1a/MODY3 gene on chromosome 12."
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS BETA AND GAMMA).
RX TISSUE=Placenta;
RL MEDLINE=97207310; PubMed=9054445;
RA Mazaki Y., Hashimoto S., Sabe H.;
RT "Monocyte cells and cancer cells express novel paxillin isoforms with

RA Bellis S.L., Miller J.T., Turner C.E.;
 RT "Characterization of tyrosine phosphorylation of paxillin in vitro by
 focal adhesion kinase.";
 RL J. Biol. Chem. 270:17437-17441(1995).
 CC -I- FUNCTION: CYTOSKELETAL PROTEIN INVOLVED IN ACTIN-MEMBRANE
 CC ATTACHMENT AT SITES OF CELL ADHESION TO THE EXTRACELLULAR MATRIX
 CC (FOCAL ADHESION). BINDS IN VITRO TO VINCULIN AS WELL AS TO THE SH3
 CC DOMAIN OF C-SRC AND, WHEN TYROSINE PHOSPHORYLATED, TO THE SH2
 CC DOMAIN OF V-CRK.
 CC -I- PTM: Phosphorylated on tyrosine residues during integrin-mediated
 CC cell adhesion, embryonic development, fibroblast transformation
 CC and following stimulation of cells by mitogens.
 CC -I- SIMILARITY: Contains 4 LIM zinc-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U14589; AAC59665.1; -;
 DR EMBL: L30099; AAC38018.1; -;
 DR PIR: B55933; B55933.
 DR HSP: P04006; LIML.
 DR InterPro: IPR001781; LIM.
 DR InterPro: IPR001904; Paxillin.
 DR Pfam: PF00412; LIM; 4.
 DR Pfam: PF03535; Paxillin; 1.
 DR PRINTS: P00832; PAXILLIN.
 DR ProDom: PD000094; LIM; 4.
 DR SMART: SM00132; LIM; 4.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 4.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 4.
 DR Cytoskeleton; Phosphorylation; LIM domain; Repeat; Metal-binding;
 KW Zinc.
 FT DOMAIN 46 53 PRO-RICH.
 FT DOMAIN 326 376 LIM 1.
 FT DOMAIN 385 435 LIM 2.
 FT DOMAIN 444 494 LIM 3.
 FT DOMAIN 503 553 LIM 4.
 FT MOD_RES 31 31 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 118 118 PHOSPHORYLATION (BY FAK1).
 SQ SEQUENCE 559 AA; 61242 MW; 6450270D90B2DE84 CRC64;

 Query Match 18.6%; Score 303.5; DB 1; Length 559;
 Best Local Similarity 27.5%; Pred. No. 7.6e-18;
 Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3;

 QY 40 CAECQQLIGHDSRELVEYDRHFEHGFRCRCQCRSLADEPFTQDSLELLCNDYCQSAFSS 99
 Db 326 CGACKKPFA--GOVVTAMGKTWHPHFVTHCQEEIGSRNFFRDCQPYCEKDYHNLFS 383
 QY 100 QCSACGETVMPGSRKLEYGGQVTHWHEHCFCLSCGCEQPLGSRFSVPDKGAHYCVCYENKFA 159
 Db 384 RCYICNGPIL--DKVVYALDRTHWPEHFFCAQCGVFFGPEGPFKDKGKAYCRKDYDFMFA 441
 QY 160 PRCARSKTLTQGGVYTRQPHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAP 219
 Db 442 PKGGCARAILENYISALNTLWHPCEFCVCRECTPFPIFNGSFFEHGDPQYCEVHYHRRGS 501
 QY 220 KCSCKRPVGLGGKGVYSPFEDRHHNCFSCARCSTSLVGCGFVDPGDQVLCQGC 275
 Db 502 LCSGCKQKPT---GRCITAMGKKFHPHFVCAFLKQLNKGTFKEQNDKPYCQNC 553

 RESULT 13
 ID_YP96_CAEEL STANDARD; PRT; 256 AA.
 AC Q09476;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical 28.9 kDa protein C28H8.6 in chromosome III.
 GN C28H8.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N2.
 RC STRAIN-Bristol N2.
 RA Miller N., Waterston R.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 CC -I- SIMILARITY: Contains 4 LIM zinc-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U20861; AAA62294.1; -;
 DR PIR: E88469; E88469.
 DR HSP: P32965; ICTL.
 DR WormPep: C28H8.6; CE01828.
 DR InterPro: IPR001781; LIM.
 DR Pfam: PF00412; LIM; 4.
 DR ProDom: PD000094; LIM; 4.
 DR SMART: SM00132; LIM; 4.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 4.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 4.
 DR Hypothetical protein; LIM domain; Metal-binding; Zinc; Repeat.
 FT DOMAIN 19 69 LIM 1.
 FT DOMAIN 78 128 LIM 2.
 FT DOMAIN 137 170 LIM 3.
 FT DOMAIN 196 246 LIM 4.
 SQ SEQUENCE 256 AA; 28901 MW; 8A6A70DB736B8225 CRC64;

 Query Match 18.4%; Score 300.5; DB 1; Length 256;
 Best Local Similarity 30.6%; Pred. No. 6.6e-18;
 Matches 57; Conservative 27; Mismatches 87; Indels 15; Gaps 3;

 QY 101 CSACGETVMPGSRKLEYGGQVTHWHEHCFCLSCGCEQPLGSRFSVPDKGAHYCVCYENKFA 160
 Db 19 CAACGKPII--GOVVTALGKWHPEHYTCCECAELGQRPFFERNGRAFCEDYHNGFSP 76
 QY 161 RCARSKTLTQGGVYTRQPHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAP 220
 Db 77 KCQCGHRAITDRCVSMVKNKFHIECTCAECNQPFGEDEGPFHEKNGQTYCKRDFRFLFAPK 136
 QY 221 CSSCKRPVGLGGKGVYSPFEDRHHNCFSCARCSTSLVGCGFVDPGDQVLCQGC----- 274
 Db 137 CNGCSQPII---SNFITALGTHWHPDCFVCHQCGVSNFGASFHEHNGAPLCERHYESR 192
 QY 275 ---CSQ 277
 Db 193 GSICSQ 198

 RESULT 14
 ID_UN97_CAEEL STANDARD; PRT; 348 AA.
 AC P50464;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE LIM protein unc-97.
 GN UNC-97 OR F14D12.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;


```
DR SMART; SM00132; LIM; 3.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS50023; LIM_DOMAIN_2; 3.
KW LIM domain; Metal-binding; Zinc; Repeat.
FT DOMAIN 184 249 LIM 1.
FT DOMAIN 250 309 LIM 2.
FT DOMAIN 310 373 LIM 3.
FT DOMAIN 150 153 POLY-GLU.
FT DOMAIN 513 522 POLY-HIS.
FT DOMAIN 544 557 POLY-SER.
FT CONFLICT 185 185 I -> S (IN REF. 3).
FT CONFLICT 370 509 MISSING (IN REF. 3).
SQ SEQUENCE 615 AA; 68609 MW; CD024365C072B052 CRC64;

Query Match 16.2%; Score 265; DB 1; Length 615;
Best Local Similarity 29.9%; Pred. No. 1.2e-14;
Matches 55; Conservative 31; Mismatches 88; Indels 10; Gaps 6;

QY 101 CSACGETVMPGSRKLEYG---GQTWHEHCFLCGCEQPLGSRSFVPDKGAHYCVPCYEN 156
Db ||: | | : | | ||: ||: | | : | | :
186 CEECGKQIGGGDIATVAFSRAGLGACWHPCFVCTTCQELLVDLIYFYHVKYVCGRRHAE 245
QY 157 KFAPRCARCSKTLTQGGVYRD-QPWHRECLVCTGCTPLAGQQTSDRDEDPYCVACFGE 215
Db ||| | : : | : | : | | : | | : | | :
246 CLRPRCQACDEIFISPECTEAEGRHWHMDHFCCFECEASLGGQRYVMRQSRPHCCACYE 305
QY 216 LFAPKCSCKRPIVGLGGKYVSFEDRHH--HNCFCARCSTSLVCGGFVDPGDQVLC- 272
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
306 RHAEYCDGCGEHI-GLDQSQ-WAYEGQHWHASDRCFCCSRGRLILGRPLPRGLIFCS 363
QY 273 QGCS 276
Db : ||
364 RACS 367
```

Search completed: August 22, 2003, 19:58:08
Job time : 26 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	1635	100.0	280	4	Q9BVA2	Q9bva2 homo sapien
2	1615	98.8	280	4	Q9P100	Q9p100 homo sapien
3	1007	61.6	172	4	Q96C98	Q96c98 homo sapien
4	991	60.6	279	11	Q8VDP9	Q8vdp9 mus muscultu
5	934	57.1	339	5	Q9VVB7	Q9vvb7 drosophila
6	934	57.1	339	5	Q8MYZ3	Q8myz3 drosophila
7	934	57.1	339	5	Q8IQ33	Q8iqz3 drosophila
8	934	57.1	339	5	Q9VVB5	Q9vvb5 drosophila
9	932	56.4	525	5	O17832	O17832 caenorhabdi
10	866	53.0	284	11	Q9WTX7	Q9wtx7 mus muscultu
11	865	52.9	284	4	Q9NTP9	Q9ntp9 homo sapien
12	851	52.7	284	4	Q9NQU2	Q9nqu2 homo sapien
13	855	52.3	284	4	Q8WZ21	Q8wz21 homo sapien
14	846	51.7	291	5	O76491	O76491 brachiosto
15	765.5	46.8	296	6	Q9GJVA	Q9gjva sus scrofa
16	760.5	46.5	280	11	Q9WUH4	Q9wuh4 rattus sp.

QY 61 FHEGCFRCRCRQSLADEPFTCDSELLCNDYCSAFSSQCSAGCTVMPGSRKLEYGG 120
 DB 61 FHEGCFRCRCRQSLADEPFTCDSELLCNDYCSAFSSQCSAGCTVMPGSRKLEYGG 120
 QY 121 TWHEHCFLCSCGEPGLGSRFVDPKGAHYCVPCYENKFAFPCARCSKTLTGGVYTRDQ 180
 DB 121 TWHEHCFLCSCGEPGLGSRFVDPKGAHYCVPCYENKFAFPCARCSKTLTGGVYTRDQ 180
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPITVGLGGKYYVSFE 240
 DB 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPITVGLGGKYYVSFE 240
 QY 241 DRHHHNCFCARCSTSLVGQGFVDPDQVLCQCSQAGP 280
 DB 241 DRHHHNCFCARCSTSLVGQGFVDPDQVLCQCSQAGP 280

RESULT 2

Q9P100 PRELIMINARY; PRT; 280 AA.
 AC Q9P100;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE LIM-only protein FHL3.
 GN FHL3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RA Lee S.M.Y., Tsui S.K.W., Chan K.K., Kotaka M., Li H.Y., Chim S.S.C.,
 RA Waye M.M.Y., Fung K.P., Lee C.Y.;
 RT "Chromosomal mapping of a skeletal muscle specific LIM-Only protein
 RT FHL3 to the distal end of the short arm of human chromosome 1.";
 RL Somat. Cell Mol. Genet. 0:0-0(1999).
 CC -1- SIMILARITY: CONTAINS 4 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 DR EMBL; AF133732; AAF61376.1; -.
 DR HSP; P32965; 1CUL.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 4.
 DR ProDom; PD000094; LIM; 4.
 DR SMART; SM00132; LIM; 4.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
 DR PROSITE; PS0023; LIM_DOMAIN_2; 4.
 KW LIM domain; Metal-binding; Zinc.
 SQ SEQUENCE 280 AA; 31282 MW; 95E437D966F7A6E7 CRC64;

Query Match 98.8%; Score 1615; DB 4; Length 280;
 Best Local Similarity 98.9%; Pred. No. 1.7e-159;
 Matches 277; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MSESFDCAKNESLYGRKIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRELFEYEDRH 60
 DB 1 MSESFDCAKNESLYGRKIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRELFEYEDRH 60
 QY 61 FHEGCFRCRCRQSLADEPFTCDSELLCNDYCSAFSSQCSAGCTVMPGSRKLEYGG 120
 DB 61 FHEGCFRCRCRQSLADEPFTCDSELLCNDYCSAFSSQCSAGCTVMPGSRKLEYGG 120
 QY 121 TWHEHCFLCSCGEPGLGSRFVDPKGAHYCVPCYENKFAFPCARCSKTLTGGVYTRDQ 180
 DB 121 TWHEHCFLCSCGEPGLGSRFVDPKGAHYCVPCYENKFAFPCARCSKTLTGGVYTRDQ 180
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPITVGLGGKYYVSFE 240
 DB 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPITVGLGGKYYVSFE 240
 QY 241 DRHHHNCFCARCSTSLVGQGFVDPDQVLCQCSQAGP 280
 DB 241 DRHHHNCFCARCSTSLVGQGFVDPDQVLCQCSQAGP 280

DB 241 DRHHHNCFCARCSTSLVGQGFVDPDQVLCQCSQAGP 280
 RESULT 3
 Q96C98 PRELIMINARY; PRT; 172 AA.
 AC Q96C98;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 3 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 DR EMBL; BC014501; AAH14501.1; -.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 3.
 DR ProDom; PD000094; LIM; 3.
 DR SMART; SM00132; LIM; 3.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE; PS0023; LIM_DOMAIN_2; 3.
 KW Hypothetical protein; LIM domain; Metal-binding; Zinc.
 SQ SEQUENCE 172 AA; 18918 MW; 5918FF5E8C015BE6 CRC64;

Query Match 61.6%; Score 1007; DB 4; Length 172;
 Best Local Similarity 100.0%; Pred. No. 1.3e-96;
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 109 MPGSRKLEYGGQTWHEHCFLCSCGCEQPLGSRFVDPKGAHYCVPCYENKFAFPCARCSKT 168
 DB 1 MPGSRKLEYGGQTWHEHCFLCSCGCEQPLGSRFVDPKGAHYCVPCYENKFAFPCARCSKT 60
 QY 169 LTQGGVYTRDQVHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPI 228
 DB 61 LTQGGVYTRDQVHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPI 120
 QY 229 VGLGGGKYVSPEDRHHHNCFCARCSTSLVGQGFVDPDQVLCQCSQAGP 280
 DB 121 VGLGGGKYVSPEDRHHHNCFCARCSTSLVGQGFVDPDQVLCQCSQAGP 172

RESULT 4

Q8VDP9 PRELIMINARY; PRT; 279 AA.
 AC Q8VDP9;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Four and a half LIM domains 2.
 GN FHL2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 4 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 DR EMBL; BC021468; AAH21468.1; -.
 DR MGD; MGI:1338762; Fhl2.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 4.
 DR ProDom; PD000094; LIM; 4.

DR SMART; SM00132; LIM; 4.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 4.
 KW LIM domain; Metal-binding; Zinc.
 SQ SEQUENCE 279 AA; 32056 MW; 11CGBC4B4131EAA7 CRC64;

Query Match 60.6%; Score 991; DB 11; Length 279;
 Best Local Similarity 54.9%; Pred. No. 9.7e-95;
 Matches 152; Conservative 57; Mismatches 68; Indels 0; Gaps 0;

QY 1 MSESFCACNESLYGRKVIQTDSPYCVPCYDNTFANTCAECQQLIGHDSRELFEDRH 60
 DB 1 MTFERDCHHNSLYGRKVIQTDSPYCVPCYDNTFANTCAECQQLIGHDSRELFEDRH 60
 QY 61 FHGECFRCRCRSLADEPFTQDSSELLCNDYCSAFSSQCSACGETVMPGSKLEYGGQ 120
 DB 61 WHGECFHCSCGSLVDKPPAAKEEQLLCTDCYSEYSSKCECKTITMPGTRMEYKGS 120
 QY 121 TWHEHFLCSCGROPGLRSFVDPDKAHYCVPCYENKFAPRCARCKTLFGGVYRDQP 180
 DB 121 SWHETCFCQRCQOPIGTGKSPFENQFCVPCYEQYALQCVQCKPKITGGVYREQP 180
 QY 181 WHRECLVCTCQPLAGQFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKYSFE 240
 DB 181 WHRECFVCTACKQLSQRTARDEPFYLCITCDLIYAKKACGTNPISLGGTKYLSFE 240
 QY 241 DRHHHNCFCARCSLSVQGGFVDPGDQVLCOGCSQ 277
 DB 241 ERQWINDCFNCKCSLSVGRGLFTRDILCPDCGK 277

RESULT 5
 Q9VVVB7 PRELIMINARY; PRT; 339 AA.
 AC Q9VVVB7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG32171 protein.
 GN LMPT OR CG9959 OR GGL1914 OR CG11916 OR CG13028 OR CG32171.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN RN
 RC SEQUENCE FROM N.A.
 RC SPRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleob J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Stadling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Matel B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleob J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 5 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 IONS.
 DR EMBL; AE003525; AAF49396.2; -
 DR HSP; P32965; ICTL.
 DR FlyBase; FBgn0036672; Lmpt.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 5.
 DR ProDom; PD000094; LIM; 5.
 DR SMART; SM00132; LIM; 5.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 5.
 KW LIM domain; Metal-binding; Zinc.
 SQ SEQUENCE 339 AA; 38804 MW; 31883C86FF6D915B CRC64;

Query Match 57.1%; Score 934; DB 5; Length 339;
 Best Local Similarity 52.2%; Pred. No. 9.7e-89;
 Matches 144; Conservative 55; Mismatches 77; Indels 0; Gaps 0;

QY 2 SESFDCACNESLYGRKVIQTDSPYCVPCYDNTFANTCAECQQLIGHDSRELFEDRH 61
 DB 60 SGHFCWCQDSELTGQRYVIRDRHPHYCIKCYENFANTCECNKIIGDSKLSYDKDHW 119
 QY 62 HEGFCRCRCRSLADEPFTQDSSELLCNDYCSAFSSQCSACGETVMPGSKLEYGGQ 121

```

Db 120 HEACFLCFKCHLSVDKQFGAKADKIYGCNGCYDAQFASRCDCGCVFVFRAGTKMEYKTRQ 179
QY 122 WHEHCFCLSCGCEQPLGSRFVDPDKGAHVCVPCYENKFPAPRCARCSKTLTGQGVYTRDOPW 181
Db 180 WHNCFCCVCKTAIGTSFTPREQEIYACGYEEKFATRCIKNKVITSGVTKNKPW 239
QY 182 HRECLVCTGCOPTLAGOQFTSRDEDPYCVACFGELFAPKCSCKRPVGLGGKVVSPED 241
Db 240 HRECTCTHCNTITLAGQFTSRDEKPYCAEFGELFAPKRCACVAPITGIGGTRFISPED 299
QY 242 RHHHNCFCARCSSTSLVGQGFVDPDQVLCQGCSCQ 277
Db 300 RHHHDCFCASCASLVGRGFTDGPDLPCDCAK 335

RESULT 6
Q8MYZ5 PRELIMINARY; PRT; 339 AA.
AC Q8MYZ5;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE RES37250p.
GN BCDNA:RE37250.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez C., Guanin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Paclele J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 5 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS
CC EMBL: AY113460; AM29465.1; -
CC FlyBase: FBgn0063025; BCDNA:RE37250.
CC InterPro: IPR001781; LIM.
CC Pfam: PF00412; LIM; 5.
CC ProDom: PD000094; LIM; 5.
CC SMART: SM00132; LIM; 5.
CC PROSITE: PS00478; LIM_DOMAIN_1; 4.
CC PROSITE: PS50023; LIM_DOMAIN_2; 5.
CC LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 339 AA; 38834 MW; AA7D2FCA9BF1FFB8 CRC64;

Query Match 57.1%; Score 934; DB 5; Length 339;
Best Local Similarity 52.2%; Pred. No. 9.7e-89;
Matches 144; Conservative 55; Mismatches 77; Indels 0; Gaps 0;

QY 2 SESFCAKNSLYGRKVIQTDSPGVCYPCVDNTFANTCAEQOOLIGHDSRELVEDRHF 61
Db 60 SHFCFCWCDSELGQVRVIRDDHPYCIKCNVENFANTCECNKIIGDSKLSKKDHW 119
QY 62 HEGCFRCRCORSLADEPFTCODSELLCNDCYSAFSSQCSACGTVMPGSKKLEYGQOT 121
Db 120 HEACFLCFKCHLSVDKQFGAKADKIYGCNGCYDAQFASRCDCGCVFVFRAGTKMEYKTRQ 179
QY 122 WHEHCFCLSCGCEQPLGSRFVDPDKGAHVCVPCYENKFPAPRCARCSKTLTGQGVYTRDOPW 181
Db 180 WHNCFCCVCKTAIGTSFTPREQEIYACGYEEKFATRCIKNKVITSGVTKNKPW 239
QY 182 HRECLVCTGCOPTLAGOQFTSRDEDPYCVACFGELFAPKCSCKRPVGLGGKVVSPED 241
Db 240 HRECTCTHCNTITLAGQFTSRDEKPYCAEFGELFAPKRCACVAPITGIGGTRFISPED 299

```

```

QY 242 RHHHNCFCARCSSTSLVGQGFVDPDQVLCQGCSCQ 277
Db 300 RHHHDCFCASCASLVGRGFTDGPDLPCDCAK 335

RESULT 7
Q8IQO3 PRELIMINARY; PRT; 558 AA.
AC Q8IQO3;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CG32171-PD.
GN CG32171.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,
RA Abail J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
RA Patuzzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez C., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Paclele J., Paragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,

```



```

CC  -!- SIMILARITY: CONTAINS 4 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC  IONS.
DR  EMBL; AL132776; CAB86657.1; -.
DR  InterPro; IPR000362; Fumarate_lyase.
DR  InterPro; IPR001781; LIM.
DR  Pfam; PF00412; LIM; 4.
DR  ProDom; PD000094; LIM; 4.
DR  SMART; SM00132; LIM; 4.
DR  PROSITE; PS00163; FUMARATE_LYASES; 1.
DR  PROSITE; PS00478; LIM_DOMAIN_1; 4.
DR  PROSITE; PS50023; LIM_DOMAIN_2; 4.
KW  LIM domain; Metal-binding; Zinc.
SQ  SEQUENCE 284 AA; 32719 MW; 82C03B1EAA147C11 CRC64;

Query Match          52.9%; Score 865; DB 4; Length 284;
Best Local Similarity 48.7%; Pred. No. 1.2e-81;
Matches 132; Conservative 52; Mismatches 87; Indels 0; Gaps 0

QY  5 FDCAKCNESLYGRXYIQTDSGYPCVPCYDNTFANTCAECQQLIGHDSRELFYEDRHFEHG 64
DB  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB  6 FYCQYCTASLLGKKVLLKDDSPYCVTCYDRVFSNVCBECKPIESDSKDLCKYKDRHWHEG 65
QY  65 CFCRCRCORSIADPEFTQDSSELLCNDYCISAFSSQCSACGETVMPGSRKLEYGGQYWH 124
DB  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB  66 CFKCTKCNHSLVERPFAAKDRLCLTECYSNECSKCFHCXRTIMPGRKMEFKNGYWH 125
QY  125 HCFLCSGEQPLGSRSPVDPKGAYHVCPCYENKFAPCARCSKTLTGCGVTVYRQDPWHRE 184
DB  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB  126 TCFVCENCRQDIPGRKPLSKESGNYVCPFEKEFAHYNCFCKKVTITSGGITPCDOLWHKE 185
QY  185 CLVCTGQOTPLAGQOFTSRDDPYCYACFGELFAPKCSSCKRPYVGLGGGYKVSFEDRHW 244
DB  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB  186 CFLCSGCKKDLCEQFMSRDDYPCVDYCNHLYANKVCACSKPISGLTGAKFICQDSQW 245
QY  245 HHNCFSCARCSTSLVGQGFVDPGQVLCQGC 275
DB  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB  246 HSECFNCGKCSVLVGKGFLLQNKEIFCQKC 276

RESULT 12
Q9NQ2 PRELIMINARY; PRT; 284 AA.
ID Q9NQ2
AC Q9NQ2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE LIM protein ACT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A..
RC TISSUE=Testis;
RX MEDLINE=21226154; PubMed=11327716;
RA Palermo I., Litrico L., Emanuele G., Gluffrida V., Sassone-Corsi P.,
RA De Cesare D., Maria Finia G., D'Agata R., Calogero A.E., Travali S.;
RT "Cloning and expression of activator of CREM in testis in human
RT testicular tissue."
RL Biochem. Biophys. Res. Commun. 283:406-411(2001).
CC -!- SIMILARITY: CONTAINS 4 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR  EMBL; AF278541; AAF85978.1; -.
DR  InterPro; IPR000362; Fumarate_lyase.
DR  InterPro; IPR001781; LIM.
DR  Pfam; PF00412; LIM; 4.
DR  ProDom; PD000094; LIM; 4.
DR  SMART; SM00132; LIM; 4.
DR  PROSITE; PS00163; FUMARATE_LYASES; 1.
DR  PROSITE; PS00478; LIM_DOMAIN_1; 4.
DR  PROSITE; PS50023; LIM_DOMAIN_2; 4.
KW LIM domain; Metal-binding; Zinc.
SQ  SEQUENCE 284 AA; 32813 MW; B37E5E9544725948 CRC64;

```

Query Match 52.7%; Score 861; DB 4; Length 284;
 Best Local Similarity 48.3%; Pred. No. 3e-81;
 Matches 131; Conservative 52; Mismatches 88; Indels 0; Gaps 0;

QY 5 FDCAKNESLYGRKVIQTDSPGVCYPCYDNTFANTCAEQQLIGHDSRELYEDRHFEH 64
 DB 6 FYCQYCTASLLGKKVYLKDDSPYCTCYDRVFSNYCECKPIESDSDKDFCYKDRHWHEG 65
 QY 65 CFRCCRCORSLADEPFTQDSELNDCYCASFSSQCSACGETVMPGSRKLEYGQOTWHE 124
 DB 66 CFKCTKCNHSLVERPFAAKDERLLCTECYSNECSKCFHCKRTIMPGRKMEFFKGNWHE 125
 QY 125 HCFLCSGCEQPLGSRFVDPDKGAHYCVPCYENKFAPCARCSKTLTQGGVTVRDQPWHE 184
 DB 126 TCFVCENCROPIGTGKPLISKESGNYCVCFEKEFAHYCNFKKVTSGITFCQDLWHE 185
 QY 185 CLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKSCSKRPVGLGGKYYVSFEDRHW 244
 DB 186 CFLCSGCRDLCEEQFMSRDDYPFCMDCYNHLYANKVACSKPISGLTGAKFICFQDSQW 245
 QY 245 HNFCSARCSTSLVGQGFVDPDGOVLQCGC 275
 DB 246 HSECFNCGKCSVSLVGKGLFTQNKEIFCQK 276

RESULT 13
 Q8WW21 PRELIMINARY; PRT; 284 AA.
 AC Q8WW21
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Similar to activator of CREM in testis (Activator of CAMP-responsive element modulator (CREM) in testis).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RA Strausberg R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 4 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC IONS.

QY 5 FDCAKNESLYGRKVIQTDSPGVCYPCYDNTFANTCAEQQLIGHDSRELYEDRHFEH 64
 DB 6 FYCQYCTASLLGKKVYLKDDSPYCTCYDRVFSNYCECKPIESDSDKDFCYKDRHWHEG 65
 QY 65 CFRCCRCORSLADEPFTQDSELNDCYCASFSSQCSACGETVMPGSRKLEYGQOTWHE 124
 DB 66 CFKCTKCNHSLVERPFAAKDERLLCTECYSNECSKCFHCKRTIMPGRKMEFFKGNWHE 125
 QY 125 HCFLCSGCEQPLGSRFVDPDKGAHYCVPCYENKFAPCARCSKTLTQGGVTVRDQPWHE 184
 DB 186 CFLCSGCRDLCEEQFMSRDDYPFCMDCYNHLYANKVACSKPISGLTGAKFICFQDSQW 245
 QY 245 HNFCSARCSTSLVGQGFVDPDGOVLQCGC 275
 DB 246 HSECFNCGKCSVSLVGKGLFTQNKEIFCQK 276

Query Match 52.3%; Score 855; DB 4; Length 284;
 Best Local Similarity 48.0%; Pred. No. 1.3e-80;
 Matches 130; Conservative 53; Mismatches 86; Indels 0; Gaps 0;

QY 5 FDCAKNESLYGRKVIQTDSPGVCYPCYDNTFANTCAEQQLIGHDSRELYEDRHFEH 64
 DB 6 FYCQYCTASLLGKKVYLKDDSPYCTCYDRVFSNYCECKPIESDSDKDFCYKDRHWHEG 65
 QY 65 CFRCCRCORSLADEPFTQDSELNDCYCASFSSQCSACGETVMPGSRKLEYGQOTWHE 124

DB 66 CFKCTKCNHSLVERPFAAKDERLLCTECYSNECSKCFHCKRTIMPGRKMEFFKGNWHE 125
 QY 125 HCFLCSGCEQPLGSRFVDPDKGAHYCVPCYENKFAPCARCSKTLTQGGVTVRDQPWHE 184
 DB 126 TCFVCENCROPIGTGKPLISKESGNYCVCFEKEFAHYCNFKKVTSGITFCQDLWHE 185
 QY 185 CLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKSCSKRPVGLGGKYYVSFEDRHW 244
 DB 186 CFLCSGCRDLCEEQFMSRDDYPFCMDCYNHLYANKVACSKPISGLTGAKFICFQDSQW 245
 QY 245 HNFCSARCSTSLVGQGFVDPDGOVLQCGC 275
 DB 246 HSECFNCGKCSVSLVGKGLFTQNKEIFCQK 276

RESULT 14
 O76491 PRELIMINARY; PRT; 291 AA.
 AC O76491
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE LIM-domain protein.
 GN AMPHIDRAL.
 OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 OX NCBI_TaxID=7739;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed-9767167;
 RA Schubert M., Holland N.D., Holland L.Z.;
 RT "Amphioxus Amphidral encoding a LIM-domain protein: Expression in the epidermis but not in the presumptive neuroectoderm."
 RL Mech. Dev. 76:203-205(1998).
 CC -1- SIMILARITY: CONTAINS 5 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC IONS.
 DR EMBL; AF071773; AAC69756.1; -.
 DR HSSP; P32965; 1CTL.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 5.
 DR ProDom; PD000094; LIM; 5.
 DR SMART; SM00132; LIM; 5.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
 DR PROSITE; PS50023; LIM_DOMAIN_2; 5.
 KW LIM domain; Metal-binding; Zinc.
 SQ SEQUENCE 291 AA; 33430 MW; 253289F21E27E4AA CRC64;

Query Match 51.7%; Score 846; DB 5; Length 291;
 Best Local Similarity 45.6%; Pred. No. 1.1e-79;
 Matches 124; Conservative 63; Mismatches 85; Indels 0; Gaps 0;

QY 5 FDCAKNESLYGRKVIQTDSPGVCYPCYDNTFANTCAEQQLIGHDSRELYEDRHFEH 64
 DB 11 FNCFCNNLSLTHGRYVNRDTHYCLKCYEKLFAPFCECHGQKIGTDVLDLDFNKKHHEQ 70
 QY 65 CFRCCRCORSLADEPFTQDSELNDCYCASFSSQCSACGETVMPGSRKLEYGQOTWHE 124
 DB 71 CFNCSKCKSLVDQOFTOKSKIKYCAQCHKEFLGKDCGCHQHPDGGKKMEYQGNWHE 130
 QY 125 HCFLCSGCEQPLGSRFVDPDKGAHYCVPCYENKFAPCARCSKTLTQGGVTVRDQPWHE 184
 DB 131 KCFTCKECKPVGTKSFIAKDDKVKTCQCYEDKVAKCKEKRKVISMGITVYKDTPWHE 190
 QY 185 CLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKSCSKRPVGLGGKYYVSFEDRHW 244
 DB 191 CFVCTHCKKPMGSRFTSKDNPNYCINCYGDLFAKCKCAKTKPTITGLGGTKFISFENSNW 250
 QY 245 HNFCSARCSTSLVGQGFVDPDGOVLQCGC 276
 DB 251 HSDCFNCTGCKTSLVGKGLFTQNKEIFCQK 282

Search completed: August 22, 2003, 19:59:56
Job time : 103 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 10:38:49 ; Search time 4833 Seconds
(without alignments)
16683.814 Million cell updates/sec

Title: US-09-786-135-2
Perfect score: 1971
Sequence: 1 ggtccgtctgcagtcgcag.....ctctgacgtggggatgaa 1971

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba:**
- 2: gb_hgt:**
- 3: gb_in:**
- 4: gb_om:**
- 5: gb_ov:**
- 6: gb_pat:**
- 7: gb_ph:**
- 8: gb_pl:**
- 9: gb_pr:**
- 10: gb_ro:**
- 11: gb_sts:**
- 12: gb_sy:**
- 13: gb_un:**
- 14: gb_vi:**
- 15: em_ba:**
- 16: em_fun:**
- 17: em_hum:**
- 18: em_in:**
- 19: em_mu:**
- 20: em_om:**
- 21: em_or:**
- 22: em_ov:**
- 23: em_pat:**
- 24: em_ph:**
- 25: em_pi:**
- 26: em_ro:**
- 27: em_sts:**
- 28: em_un:**
- 29: em_vi:**
- 30: em_htg_hum:**
- 31: em_htg_inv:**
- 32: em_htg_other:**
- 33: em_htg_mus:**
- 34: em_htg_pln:**
- 35: em_htg_rod:**
- 36: em_htg_man:**
- 37: em_htg_vrt:**
- 38: em_sy:**
- 39: em_htgo_hum:**
- 40: em_htgo_mus:**
- 41: em_htgo_other:**

Pred. No. is the number of results, predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1514.2	76.8	1690	9	BC001351	BC001351 Homo sapi
2	1514.2	76.8	1691	9	BC011697	BC011697 Homo sapi
3	1473.6	74.8	1497	9	AF133732	AF133732 Homo sapi
4	1466.6	74.4	1516	9	HSU60116	U60116 Homo sapien
5	1328.2	67.4	1492	9	BC014501	BC014501 Homo sapi
c	915.2	46.4	99687	9	AL603790	AL603790 Human DNA
6	871.6	44.2	1146	10	AF134772	AF134772 Mus muscu
7	842	42.7	843	9	BT007052	BT007052 Homo sapi
8	841	42.7	843	12	BT007804	BT007804 Synthetic
9	807.6	41.0	997	4	AY277587	AY277587 Sus scrof
10	711	36.1	1147	10	AF114382	AF114382 Mus muscu
11	541.4	27.5	656	10	AF149826	AF149826 Mus muscu
c	402.2	20.4	241886	10	AC027285	AC027285 Mus muscu
13	344	17.5	840	10	AB008571	AB008571 Rattus no
14	333.4	16.9	876	10	AF114381	AF114381 Mus muscu
15	333.4	16.9	1300	10	AF055889	AF055889 Mus muscu
16	332.4	16.9	1433	9	HUMDRAL	L42176 Homo sapien
17	332.4	16.9	1585	6	AX015392	AX015392 Sequence
18	332.4	16.9	1892	9	BC014397	BC014397 Homo sapi
19	331.2	16.8	840	9	BT006960	BT006960 Homo sapi
20	331.2	16.8	840	12	BT007658	BT007658 Synthetic
21	330.8	16.8	1416	9	HSU29332	U29332 Homo sapien
22	326.8	16.6	263370	2	AC122957	AC122957 Rattus no
c	326.8	16.6	279467	2	AC112814	AC112814 Rattus no
24	323.8	16.4	1378	10	AF153340	AF153340 Mus muscu
25	323.8	16.4	1450	10	MMU77040	U77040 Mus muscu
26	323.8	16.4	1487	10	BC021468	BC021468 Mus muscu
27	323.8	16.4	1487	10	BC021468	BC021468 Mus muscu
c	306.2	15.5	208285	10	AC0606907	AL606907 Mouse DNA
29	306.2	15.5	219825	10	AC098886	AC098886 Mus muscu
30	296.4	15.0	318	6	AX210478	AX210478 Sequence
31	280.4	14.2	2579	4	SSC275968	AJ275968 Sus scrof
c	280.2	14.2	108296	2	AC112815	Continuation (4 of
32	276.2	14.0	1078	10	AF083394	AF083394 Mus muscu
33	276.2	14.0	1136	10	BC048473	BC048473 Mus muscu
34	265.4	13.5	2913	3	AY113460	AY113460 Drosophil
35	265.4	13.5	2916	3	AY069696	AY069696 Drosophil
36	265.4	13.5	2916	3	DME418713	AJ418713 Drosophil
37	265.4	13.3	1320	9	AK058175	AK058175 Homo sapi
38	262.4	13.3	1320	9	AK058175	AK058175 Homo sapi
39	262.4	13.3	1627	9	AF278541	AF278541 Homo sapi
40	262.4	13.3	1846	9	AK057611	AK057611 Homo sapi
41	260.8	13.2	1089	9	BC021723	BC021723 Homo sapi
42	260.8	13.2	1321	9	BC029490	BC029490 Homo sapi
43	257.2	13.0	2592	3	AF071773	AF071773 Branchios
44	256	13.0	1200	10	AF134773	AF134773 Rattus sp
45	255.4	13.0	966	9	HSU29538	U29538 Human heart

ALIGNMENTS

RESULT 1
BC001351
LOCUS BC001351 1690 bp mRNA linear PRI 06-JUN-2003
DEFINITION Homo sapiens cDNA clone MGC:8696 IMAGE:2964682, complete cds.
ACCESSION BC001351
VERSION BC001351.1 GI:12655006
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1690)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.E., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouford, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalhus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

2 (bases 1 to 1690)

Strausberg, R.

Direct Submission

Submitted (12-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgpbbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Iaric, P., Legaspi, R., Maduro, O.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgone, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 1 Row: p Column: 15

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4758373.

FEATURES

source

1..1690

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:8696 IMAGE:2964682"

/tissue_type="Muscle, rhabdomyosarcoma"

/clone_lib="NIH_MGC_17"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

179..1021

/codon_start=1

/product="Unknown (protein for MGC:8696)"

/protein_id="AAH01351.1"

/db_xref="GI:12655007"

/translation="MSESFDCAKNCSLYGRKYIQTDSGPCVPCYDNTFANTCAEQQLIHDSRELFEYDRHFEGCFRCRCQSLADEPFTQDSELLNCDCYCSAFSSQCSACGVTMPGSRKLEGGOTWHECHFLCSCGCPGSRFVDPKGAHYCPVYENKFAFPCARCSKLTGGGVTRDQPHWRHRECLVCTGGOTPLAGOOFTSRDEDPYCVACFGELFAKCSCKRPIVLGGGKYVSPEDRHHNCFSCARCSLSLVGQGFVDPDQVLCOGCSQAGP"

CDS

BASE COUNT	336 a	550 c	452 g	352 t	
ORIGIN					
Query Match	76.8%; Score 1514.2; DB 9; Length 1690;				
Best Local Similarity	98.2%; Pred. No. 0; Mismatches 28; Indels 1; Gaps 1;				
Matches 1542; Conservative	0;				
QY	11	CAGTCGGGACGCTCGCGGACGCTCGCTCGGCGCCCGCTGTGGC-CGCGCCCGCAGGGTTC	69		
Db	104	CGGTTCCCGCCCGCGCTCGCGCGCGCTCCCGCTCCCGCGCGCTTCGAGGGTTC	163		
QY	70	TCTCCCTTGGCCACCATGAGCGAGTCAATTGACTGTGCAAAATGCAACGAGTCCCTGTAT	129		
Db	164	TCTCCCTTGGCCACCATGAGCGAGTCAATTGACTGTGCAAAATGCAACGAGTCCCTGTAT	223		
QY	130	GGACCGAAGTACATCCAGACAGACGCGCCCTACTGTGTCCTCTGATGACATACC	189		
Db	224	GGACCGAAGTACATCCAGACAGACGCGCCCTACTGTGTCCTCTGATGACATACC	283		
QY	190	TTTGGCAACACCTGTGCTGAGTGCCAGCAGCTTATCGGGCATGACTCGAGGAGGTTC	249		
Db	284	TTTGGCAACACCTGTGCTGAGTGCCAGCAGCTTATCGGGCATGACTCGAGGAGGTTC	343		
QY	250	TATGAAGACCGCCATTTCCACGAGGCTGCTTCCCGTGTGCGGTGCGCAGCGCTCACTA	309		
Db	344	TATGAAGACCGCCATTTCCACGAGGCTGCTTCCCGTGTGCGGTGCGCAGCGCTCACTA	403		
QY	310	GCCGATGAACCTTCCACCTGCCAGCAGTGAAGTGTCTGCAANTGACTGCTACTGCACT	369		
Db	404	GCCGATGAACCTTCCACCTGCCAGCAGTGAAGTGTCTGCAANTGACTGCTACTGCACT	463		
QY	370	GGGTTTCTCGCAGTGTCTCCCTTGTGGGAGACTGTCTATGCTGGGTCCCGGAAGCTG	429		
Db	464	GGGTTTCTCGCAGTGTCTCCCTTGTGGGAGACTGTCTATGCTGGGTCCCGGAAGCTG	523		
QY	430	GAATATGAGGCGCAGCATGGCATGAGCACTGCTTCTGTGTGAGTGGGTGTGAACAGCCA	489		
Db	524	GAATATGAGGCGCAGCATGGCATGAGCACTGCTTCTGTGTGAGTGGGTGTGAACAGCCA	583		
QY	490	CTGGGCTCCGCTCTTTGTGCTCCGACAGGCTGCTACTACTGCTGCTGCTGCTATGAG	549		
Db	584	CTGGGCTCCGCTCTTTGTGCTCCGACAGGCTGCTACTACTGCTGCTGCTGCTATGAG	643		
QY	550	AACAAATTTGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	609		
Db	644	AACAAATTTGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	703		
QY	610	TACCTGTATCAGCCCTGGCATCGAATGTCTGTGCTGTACCGGATGCGAGACGCCCTGTG	669		
Db	704	TACCTGTATCAGCCCTGGCATCGAATGTCTGTGCTGTACCGGATGCGAGACGCCCTGTG	763		
QY	670	GCAGGCGACGAGTTCACCTCCCGGATGAAGATCCCTACTGTGTGCTGCTGCTGCTGCTG	729		
Db	764	GCAGGCGACGAGTTCACCTCCCGGATGAAGATCCCTACTGTGTGCTGCTGCTGCTGCTG	823		
QY	730	CTCTTTGCACTAAGTGCAGCAGCTGCAAGCGCCCATCTGTAGGACTCGGTGGAGGCAAG	789		
Db	824	CTCTTTGCACTAAGTGCAGCAGCTGCAAGCGCCCATCTGTAGGACTCGGTGGAGGCAAG	883		
QY	790	TATGTGCTCTTTGAAGACCGGACACTGGCACCACAACTGCTTCTTCTGCGCCGCTGCTCT	849		
Db	884	TATGTGCTCTTTGAAGACCGGACACTGGCACCACAACTGCTTCTTCTGCGCCGCTGCTCT	943		
QY	850	ACCTCCCTGTTGGGCGCAGGGCTTCTGTACCGGATGAGACCAAGTGTCTGTGCGAGGGCTGT	909		
Db	944	ACCTCCCTGTTGGGCGCAGGGCTTCTGTACCGGATGAGACCAAGTGTCTGTGCGAGGGCTGT	1003		
QY	910	AGCCAGGCGGCGCTTAAGCCAGGGCTCTTGGACCCAGGCTTTCATACACACGCGGCCA	969		
Db	1004	AGCCAGGCGGCGCTTAAGCCAGGGCTCTTGGACCCAGGCTTTCATACACACGCGGCCA	1063		
QY	970	GGACTGTGGCTCTTTTCTAAACCACTCTGGGACTCAGCTCCCCCGCCCAAAAAATG	1029		

```

Db 1064 GGAAGTGGGCTCCTTTCTAAACACACCTCTGGGACTCAGCTCCCGCCGCAAAAAATG 1123
Qy 1030 GGTCTCCTCTGGGCTCCAGGATTTGTCCGCCACTCCAGCATCCCAAACTGTACTCCC 1089
Db 1124 GGTCTCCTCTGGGCTCCAGGATTTGTCCGCCACTCCAGCATCCCAAACTGTACTCCC 1183
Qy 1090 TGACCCAGGCCCCAATCTCTGGGCTCTTACAGAGCTTCCATGAGTCAAGCCCCCTCCCA 1149
Db 1184 TGACCCAGGCCCCAATCTCTGGGCTCTTACAGAGCTTCCATGAGTCAAGCCCCCTCCCA 1243
Qy 1150 CACTGTGACATCCAGATTCACCTCTCCCTCGAGTCTGGTTCACAGATGAGTCTCT 1209
Db 1244 CACTGTGACATCCAGATTCACCTCTCCCTCGAGTCTGGTTCACAGATGAGTCTCT 1303
Qy 1210 CCCAAATCAGGCTCTAGACCGAGCCCTCCAAACCTGGACTCTGGGACTTAGGCCCC 1269
Db 1304 CCCAAATCAGGCTCTAGACCGAGCCCTCCAAACCTGGACTCTGGGACTTAGGCCCC 1363
Qy 1270 TTAATCTAGACTCTCTTATAGGTTTCAGGTTCTCTATGGTTCCTATGGTTCCTTG 1329
Db 1364 TTAATCTAGACTCTCTTATAGGTTTCAGGTTCTCTATGGTTCCTATGGTTCCTTG 1423
Qy 1330 AAAGTGGACTGTCTCAGGCTTGACCTGCCCGCCACCCCATCCCGCGTTGAGGCTGTGG 1389
Db 1424 AAAGTGGACTGTCTCAGGCTTGACCTGCCCGCCACCCCATCCCGCGTTGAGGCTGTGG 1483
Qy 1390 GGCAGCAGATCAGGAGCCCACTGATAGGGGCCCTAGGTTACAGGTTGCTGCCAGCAGG 1449
Db 1484 GGCAGCAGATCAGGAGCCCACTGATAGGGGCCCTAGGTTACAGGTTGCTGCCAGCAGG 1543
Qy 1450 TCGCCACCGAGTGTCTCTCATTTTATTTTTCAGCTTCCATTTTGGCCATAGATGGCAGG 1509
Db 1544 TCGCCACCGAGTGTCTCTCATTTTATTTTTCAGCTTCCATTTTGGCCATAGATGGCAGG 1603
Qy 1510 GGTGAGATTGGCTCATCCCTTCCAGATTTCTCAATAAGCGGTGTGAGAGCAGGGA 1569
Db 1604 GGTGAGATTGGCTCATCCCTTCCAGATTTCTCAATAAGCGGTGTGAGAGAAAAA 1663
Qy 1570 AAAAAA 1580
Db 1664 AAAAAA 1674

RESULT 2
BC011697 1691 bp mRNA linear PRI 02-AUG-2001
LOCUS Homo sapiens, clone MGC:19547 IMAGE:3532981, mRNA, complete cds.
DEFINITION BC011697.1 GI:15079779
ACCESSION MGC.
VERSION Homo sapiens (human)
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

```

Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tlonsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 27 Row: d Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4758373.

FEATURES

source

```

1. 1691
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="MGC:19547 IMAGE:3532981"
  /tissue_type="Muscle, rhabdomyosarcoma"
  /clone_lib="NIH_MGC_17"
  /lab_host="DH10B-R"
  /note="Vector: pOTB7"
  179..1021
  /codon_start=1
  /product="unknown (protein for MGC:19547)"
  /protein_id="AAH11697.1"
  /db_xref="GI:15079780"
  /translation="MSESPDCAKNESLYGRKYIQTDSGPYCPVPCYDNTANTCAEQ
  QLGHDSRELFEDRHFGECRCRQSRSLADEPFTQDSSELNLCYSAFSSQCS
  ACGETWPMGRKLEYGGOTWHEHFLCSCGCEPLGSRSPVPCYKHAHVCYFENKAP
  RCARCKTLTQGGVTYVTPWHRLETCGTCTPLAQGQTSRDDETCYVACFGELEFA
  PKGSSCKRPITVGLGGKYVSFEDRHWHNCFSCARCSTSLVGQFVDPQVLCQGS
  QAGP"

```

CDS

```

BASE COUNT 337 a 550 c 452 g 352 t
ORIGIN
Query Match 76.8%; Score 1514.2; DB 9; Length 1691;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1542; Conservative 0; Mismatches 28; Indels 1; Gaps 1;
Qy 11 CAGTCGGCAGCTCGCCGGCAGCTCGCTCGGCCCGCCCTGTCGCTCCCGCCGCTTCGAGGGTTC 69
Db 104 CGGTTCCCGCCCGCCCTCGCGCGCTCCCGCTCCCGCCGCTTCGAGGGTTC 163
Qy 70 TCTCCCTTCCCAACATGAGGAGTCAATTTGACTGTGCAAAATGCAACAGTCTCTGAT 129
Db 164 TCTCCCTTCCCAACATGAGGAGTCAATTTGACTGTGCAAAATGCAACAGTCTCTGAT 223
Qy 130 GGACGCAAGTACATCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTATGACAAATCC 189
Db 224 GGACGCAAGTACATCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTATGACAAATCC 283
Qy 190 TTTGCCAAACCTGTGTGAGTGCCACAGCTTATCGGGATGATGACTCGAGGAGCTGTC 249
Db 284 TTTGCCAAACCTGTGTGAGTGCCACAGCTTATCGGGATGATGACTCGAGGAGCTGTC 343
Qy 250 TATGACAGCCGCTATTCACAGAGGCTGCTTCGGCTGCTGCCCTGCTCCAGCGCTCACTA 309
Db 344 TATGACAGCCGCTATTCACAGAGGCTGCTTCGGCTGCTGCCCTGCTCCAGCGCTCACTA 403
Qy 310 GCCGATGAACCTTCACTGCCAGGACAGTGTCTGCTGCAATGACTGCTACTGAGT 369
Db 404 GCCGATGAACCTTCACTGCCAGGACAGTGTCTGCTGCAATGACTGCTACTGAGT 463
Qy 370 CGGTTTCTCGCAGTGTCTCGCTTGTGGGAGACTGTCTGCTGCTGCTGCTGCTGCTGCTG 429
Db 464 CGGTTTCTCGCAGTGTCTCGCTTGTGGGAGACTGTCTGCTGCTGCTGCTGCTGCTGCTG 523
Qy 430 GAATATGAGGAGCCAGACATGACATGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 489
Db 524 GAATATGAGGAGCCAGACATGACATGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 583
Qy 490 CTGGGCTCCCGCTTCTTTTGTGCGCCGACAGGGTGTCTCACTACTGCTGCTGCTGCTGCTG 549

```

```
Db 584 CTGGGCTCCGTTCTTTTGTGCGGACAAAGGTTGCTCACTACTGCTGCCCTGCTATGAG 643
QY 550 ACAAAGTTTCTCTCGCTGCGCCGCTGCACGACAGACGCTGACACAGGTTGAGTGACA 609
Db 644 ACAAGTTTCTCTCGCTGCGCCGCTGCACGACAGACGCTGACACAGGTTGAGTGACA 703
QY 610 TACCGTGATCAGCCGTTGCGATGAGAAATGCTGGTGTGTACCGGATGCCAGACGCCCTG 669
Db 704 TACCGTGATCAGCCGTTGCGATGAGAAATGCTGGTGTGTACCGGATGCCAGACGCCCTG 763
QY 670 GCAGGCGACAGTTACCTCCCGGGATGAAGATCCCTACTGTGTGGCCCTGTTTTGAGAA 729
Db 764 GCAGGCGACAGTTACCTCCCGGGATGAAGATCCCTACTGTGTGGCCCTGTTTTGAGAA 823
QY 730 CTTCTTTCACCTTAAGTGACAGCAGCTCAAGCGCCCATCGTAGGACTCGGTGGAGGCAAG 789
Db 824 CTTCTTTCACCTTAAGTGACAGCAGCTCAAGCGCCCATCGTAGGACTCGGTGGAGGCAAG 883
QY 790 TATGTGTCTCTTTGAAGACCGACACTGGCACCACTGCTTCTCTGCGCCGCTGCTCT 849
Db 884 TATGTGTCTCTTTGAAGACCGACACTGGCACCACTGCTTCTCTGCGCCGCTGCTCT 943
QY 850 ACTCTCTGTTGGGCGCAGGCTTGTACCGGATGGAGACCAAGTGCTGTGCGAGGCTGT 909
Db 944 ACTCTCTGTTGGGCGCAGGCTTGTACCGGATGGAGACCAAGTGCTGTGCGAGGCTGT 1003
QY 910 AGCCAGGCGAGGCTTCAAGCAGGCTCTGTCGACCCAGGCTTTCCTATACACGCGCCA 969
Db 1004 AGCCAGGCGAGGCTTCAAGCAGGCTCTGTCGACCCAGGCTTTCCTATACACGCGCCA 1063
QY 970 GGACTGTGCTCTCTTTTAAACCACTCTGGGACTCAGCTCCCGCCGCCCAAAAAATG 1029
Db 1064 GGACTGTGCTCTCTTTTAAACCACTCTGGGACTCAGCTCCCGCCGCCCAAAAAATG 1123
QY 1030 GGTCTCTCTCTGGGCTCCAGGATGTCCTCCCACTCCAGCATCCCAAACTGGTACTCC 1089
Db 1124 GGTCTCTCTCTGGGCTCCAGGATGTCCTCCCACTCCAGCATCCCAAACTGGTACTCC 1183
QY 1090 TGACCCAGGCGCCCAATCTCTGGGCTTTACAGAGGCTTCCATGAGTCAAGCCCTCCCA 1149
Db 1184 TGACCCAGGCGCCCAATCTCTGGGCTTTACAGAGGCTTCCATGAGTCAAGCCCTCCCA 1243
QY 1150 CACTGTGACTCCAGAAATACCTCTCCGCTGCAGTCTGGTTCCTCAGACTGAGTCTCT 1209
Db 1244 CACTGTGACTCCAGAAATACCTCTCCGCTGCAGTCTGGTTCCTCAGACTGAGTCTCT 1303
QY 1210 CCCCAGGCTGCTAGACCCGAGCCCTCCAAACCTGGACTCTGGGACTTAGGCCCCC 1269
Db 1304 CCCCAGGCTGCTAGACCCGAGCCCTCCAAACCTGGACTCTGGGACTTAGGCCCCC 1363
QY 1270 TTAATCTAGACTCTCTTTTATAGGTTTTCAGGCTCTCTATGGTGGCTGGGAAGTCTTG 1329
Db 1364 TTAATCTAGACTCTCTTTTATAGGTTTTCAGGCTCTCTATGGTGGCTGGGAAGTCTTG 1423
QY 1330 AAAGTGGACTGTTCTCAGGCTTACCTGCCACCCCATCCCGCGGTTGAGGCTGTGG 1389
Db 1424 AAAGTGGACTGTTCTCAGGCTTACCTGCCACCCCATCCCGCGGTTGAGGCTGTGG 1483
QY 1390 GGCAGCATCAGAGGCGCCACTGATAAGGGGCGCTTAGGGTACAGGGTGTGCCAGCAGG 1449
Db 1484 GGCAGCATCAGAGGCGCCACTGATAAGGGGCGCTTAGGGTACAGGGTGTGCCAGCAGG 1543
QY 1450 TGGCCACCGAGTGTCTCTCTATTTTATTTTACGCTCCATTTTCCCATAGATGGGAGAGG 1509
Db 1544 TGGCCACCGAGTGTCTCTCTATTTTATTTTACGCTCCATTTTCCCATAGATGGGAGAGG 1603
QY 1510 GGTGAGATGGCTATCCCGCTTCCAGATTTCTGCAATAAAGCGGTGTGAGGAAGCAGGGA 1569
Db 1604 GGTGAGATGGCTATCCCGCTTCCAGATTTCTGCAATAAAGCGGTGTGAGGAAGCAGGGA 1663
QY 1570 AAAAAA 1580
|||||
```

```
Db 1664 AAAAAA 1674
RESULT 3
AF133732
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1497
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="lp34-p35"
/tissue_type="skeletal muscle"
1..1497
/gene="PHL3"
/feature="PHL3"
/feature="Four and a half LIM domains 3"
1..843
/gene="PHL3"
/codon_start=1
/product="LIM-only protein PHL3"
/protein_id="AAF61376.1"
/db_xref="GI:7381058"
/translation="MSEFDCAKNESLYGRKYIQDPSGVCPCYDNTFANTCAECQ
OLIGHDSRELVEFHEGFCRCRCORSLADEFTQDSLELLNCDCYCSAFSQC
ACGETVMPGSRKLEVGOTGHEHCFLCSCQPLGSRFPKDGHYCPVCYENKFA
SCARSKTLTGQGVTRQDPWRECLVCTGGTPLARQOFTSRDEDPYCVAGFLEA
PKSSCKRPIVLGGGKYVSFEDRHHNCFSCARCSISLVGGFVPGDQVLCGCF
QAGP"
BASE COUNT 309 a 465 c 391 g 332 t
ORIGIN
Query Match 74.8%; Score 1473.6; DB 9; Length 1497;
Best Local Similarity 99.1%; Pred. No. 0; Mismatches 14; Indels 0; Gaps 0;
Matches 1482; Conservative 0;
QY 85 ATGACGAGTCAATTGACTGTGCAAAATGCAACGAGTCCCTGTATGAGCGCAAGTACATC 144
Db 1 ATGACGAGTCAATTGACTGTGCAAAATGCAACGAGTCCCTGTATGAGCGCAAGTACATC 60
QY 145 CAGACAGACAGGGCCCCCTACTGTGTGCCCTGCTATGACAATACCTTTGCCAACACCTGT 204
Db 61 CAGACAGACAGGGCCCCCTACTGTGTGCCCTGCTATGACAATACCTTTGCCAACACCTGT 120
QY 205 GCTGAGTCCGACAGCTTATCGGGCATGACTCGAGGAGCTGCTTCTATCAAGACCGCCAT 264
Db 121 GCTGAGTCCGACAGCTTATCGGGCATGACTCGAGGAGCTGCTTCTATGAAGACCGCCAT 180
QY 265 TTCCACAGAGGCTGCTTCCGCTGCTGCCGCTGCCAGCGCTCACTACCGCATGAACCCCTTC 324
Db 181 TTCCACAGAGGCTGCTTCCGCTGCTGCCGCTGCCAGCGCTCACTACCGCATGAACCCCTTC 240
QY 325 ACCTGCCAGGACAGTGAAGTGTCTCTGCAATGACTGCTACTGCAAGTGGCGTTCCTTCGCGAG 384
```


misc_feature		60..132		/note="encodes HL-domain"	
misc_feature		157..315		/note="encodes LIM-domain 1"	
misc_feature		340..498		/note="encodes LIM-domain 2"	
misc_feature		523..675		/note="encodes LIM-domain 3"	
misc_feature		700..864		/note="encodes LIM-domain 4"	
polyA_signal		1499..1504			
polyA_site		1516			
BASE COUNT		300 a	480 c	393 g	341 t
ORIGIN		2 others			
Query Match					
Best Local Similarity 74.4%; Score 1466.6; DB 9; Length 1516;					
Matches 1484; Conservative 0; Mismatches 31; Indels 0; Gaps 0;					
QY	47	GCTTGGCCCGCCGAGGGTTCTCTCCCTTGGCCACCATGAGGAGTCAATTGACTGTG	106		
Db	2	GCACGAGCCGNTTCGAGGGTTCTCTCCCTTGGCCACCATGAGGAGTCAATTGACTGTG	61		
QY	107	CAAAATGCAAGAGTCCCTGTATGGAGCGCAAGTACATCCACAGACAGACGCGCCCTACT	166		
Db	62	CAAAATGCAAGAGTCCCTGTATGGAGCGCAAGTACATCCACAGACAGACGCGCCCTACT	121		
QY	167	GTGTGCCCTGCTATGACAATACCTTTGCCAACACTGCTGCTGAGTGCCAGCAGCTTATCG	226		
Db	122	GTGTGCCCTGCTATGACAATACCTTTGCCAACACTGCTGCTGAGTGCCAGCAGCTTATCG	181		
QY	227	GGCATGACTCGAGGAGCTGTTCCTATGAAGACCGCCATTTCCAGAGGGGTGCTTCGGCT	286		
Db	182	GGCATGACTCGAGGAGCTGTTCCTATGAAGACCGCCATTTCCAGAGGGGTGCTTCGGCT	241		
QY	287	GCTGCCGTGCGAGCGCTCACTAGCCGATGAACCTTTCACTGCCAGGACAGTGAAGTGC	346		
Db	242	GCTGCCGTGCGAGCGCTCACTAGCCGATGAACCTTTCACTGCCAGGACAGTGAAGTGC	301		
QY	347	TCTGCAATGACTGCTACTGCAGTGGTTCCTCGCAGTGTCTCGCTTGTGGGAGACTG	406		
Db	302	TCTGCAATGACTGCTACTGCAGTGGTTCCTCGCAGTGTCTCGCTTGTGGGAGACTG	361		
QY	407	TCATGCCCTGGTCCCGGAAGCTGGAATATGGAGGCCACAGATGCGATGAGCACTGCTTCC	466		
Db	362	TCATGCCCTGGTCCCGGAAGCTGGAATATGGAGGCCAAACATGCGATGAGCACTGCTTCC	421		
QY	467	TGTGCAATGGCTGTGAACAGCCACTGGGCTCCCGTCTTTTGTGCCGCAAGGGTGTCT	526		
Db	422	TGTGCAATGGCTGTGAACAGCCACTGGGCTCCCGTCTTTTGTGCCGCAAGGGTGTCT	481		
QY	527	ACTACTGCGTGGCTGTATGAGAACAAAGTTTGTCTCGTGGCGCGCTGACGCAAGA	586		
Db	482	ACTACTGCGTGGCTGTATGAGAACAAATTTGTCTCGTGGCGCGCTGACCAAAA	541		
QY	587	CGCTGACACAGGTGGAGTGACATACCGTGATCAGCGTGCGATGAGCAATGTCTGCTCT	646		
Db	542	CGCTTACACAGGTGGATTGACATACCGTGATCTGCCGTGGCATCCAAAATGTCTGCTCT	601		
QY	647	GTACCGGATGCCAGACGCCCTTGGCAGGGCAGAGTTTCACTTCCCGGGATGAAGATCCCT	706		
Db	602	GTACCGGATGCCAGACGCCCTTGGCAGGGCAGAGTTTCACTTCCCGGGATGAAGATCCCT	661		
QY	707	ACTGTGTGGCTGTTTGGAGAACTCTTTGCACCTTAAGTCAGCAGCTGCAAGGCCCA	766		
Db	662	ACTGTGTGGCTGTTTGGAGAACTCTTTGCACCTTAAGTCAGCAGCTGCAAGGCCCA	721		
QY	767	TCGTAGGACTCGGTGGAGGCAAGTATGTCTTTGAAGCCGACACTGGCACCACAACT	826		
Db	722	TCGTAGGACTCGGTGGAGGCAAGTATGTCTTTGAAGCCGACACTGGCACCACAACT	781		
QY	827	GCTTCTCTCGCCCGCTGCTCTACCTCCCTTGTGGGCCAGGGCTTCTGACCGGATGGAG	886		

Db	782	GCTTCACTGCGACCGCTGCTCTAACTCCCTGGTGGGCCAGGGCTCTGTACCGGATGGAG	841	
QY	887	ACCAAGTCTCTGCCAGGGCTGTAGCCAGGAGGCGCCCTAAGCCAGGGCTCTCTGGACCA	946	
Db	842	ACCAAGTCTCTGCCAGGGCTGTAGCCAGGAGGCGCCCTAAGCCAGGGCTCTCTGGACCA	901	
QY	947	GGCTTTCCCATPACCACGCGGCCAGGACTGTGGCTCTCTTTCTAAACCACTCTGGGACTC	1006	
Db	902	GGCTTTCCCATPACCACGCGGCCAGGACTGTGGCTCTCTTTCTAAACCACTCTGGGACTC	961	
QY	1007	AGCTCCCGCCGCAAAAAAATGGGTCTCTCTTGGGCTCAGGATGTCTCCCCACATCC	1066	
Db	962	AGCTCCCGCCGCAAAAAAATGGGTCTCTCTTGGGCTCAGGATGTCTCCCCACATCC	1021	
QY	1067	AGCATCCCAAACTGGTACTCCCTGACCCAGGCGCCCAATCTCTGGGCTCTTACAGAGCT	1126	
Db	1022	AGCATCCCAAACTGGTACTCCCTGACCCAGGCGCCCAATCTCTGGGCTCTTACAGAGCT	1081	
QY	1127	CCATGAGTCAAGCCCGCTCCCAACACTGGACTCCAGAAATTCACCCCTCTCCCGTGCATC	1186	
Db	1082	CCATGAGTCAAGCCCGCTCCCAACACTGGACTCCAGAAATTCACCCCTCTCCCGTGCATC	1141	
QY	1187	TGGGTTCCACAGACTGAGTCTCTCCCAAAATCAGGGCTCTAGACCCGAGCCCTCAAAAC	1246	
Db	1142	TGGGTTCCACAGACTGAGTCTCTCCCAAAATCAGGGCTCTAGACCCGAGCCCTCAAAAC	1201	
QY	1247	TGGACTCTGGGACTTAGGCGCCCTTAAATCTAGACTCTCTTTATAGCTTTTTCAGGTCTCC	1306	
Db	1202	TGGACTCTGGGACTTAGGCGCCCTTAAATCTAGACTCTCTTTATAGCTTTTTCAGGTCTCC	1261	
QY	1307	TATGGGTGCTGGGAGTCTTGAAGTGAGTGTCTCTAGGCTGTGCTGACCTGCCCCACCCC	1366	
Db	1262	TATGGGTGCTGGGAGTCTTGAAGTGAGTGTCTCTGACCTGACCTGCCCCACCCC	1321	
QY	1367	ATCCCCGGCTTGGGCTGTGGGGCAGCAGATCAGGAGCCACTGATGAAGGGGCCCTAG	1426	
Db	1322	ATCCCCGGCTTGGGCTGTGGGGCAGCAGATCAGGAGCCACTGATGAAGGGGCCCTAG	1381	
QY	1427	GGTACAGGGTGTGCTGCCAGAGTGTGCTGCTGCTCTCTCATTTTATTTTTCAGTCTCA	1486	
Db	1382	GGTACAGGGTGTGCTGCCAGAGTGTGCTGCTGCTCTCTCATTTTATTTTTCAGTCTCA	1441	
QY	1487	TTTTTGGCCATAGATGGGCGAGGGTGTGATGGCTCATCCCCCTTCAGATTTCTGCAAT	1546	
Db	1442	TTTTTGGCCATAGATGGGCGAGGGTGTGATGGCTCATCCCCCTTCAGATTTCTGCAAT	1501	
QY	1547	AAAGCGGTGTGAGGA 1561		
Db	1502	AAAGCGGTGTGAGGA 1516		
RESULT 5				
BC014501				
LOCUS				
DEFINITION Homo sapiens, clone MGC:23614 IMAGE:4860329, mRNA, complete cds.				
ACCESSION BC014501				
VERSION BC014501.1 GI:15680276				
KEYWORDS MGC.				
SOURCE Homo sapiens (human)				
ORGANISM Homo sapiens				
REFERENCE 1 (bases 1 to 1492)				
AUTHORS Strausberg, R.				
TITLE Direct Submission				
JOURNAL Submitted (17-SEP-2001) National Institutes of Health, Mammalian				
Gene Collection (MGC), Cancer Genomics Office, National Cancer				
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
USA				
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT Contact: MGC help desk				
Email: cgapds-r@mail.nih.gov				
Tissue Procurement: ATCC/DCTD/DTP				

cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@gscc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 34 Row: f Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES

Location/Qualifiers

1..1492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:23614 IMAGE:4860329"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="NH.MGC.49"
/lab_host="DH10B-R"
/note="vector: pOTB"
315..833
/codon_start=1
/product="Unknown (protein for MGC:23614)"
/protein_id="AAH14501.1"
/db_xref="GI:1568027"

CDS

/translation="MPSRKLEYGGGTWHEHFLCSGCEPGLSRSFVDPKGAHYCPV
CYBNKAPARCSTLTGTGQVYRDPWHRECLVCTGCTPLAGQOFTSRDEPYCV
ACFGEIAPKCSCKRPVIGLGGKYVSFEDRHHNCFSCARCSLTVQGGFVPDGD
QVLCQSGQAGP"
BASE COUNT 284 a 491 c 403 g 314 t
ORIGIN

Query Match 67.4%; Score 1328.2; DB 9; Length 1492;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 240 GGAGCTCTTATGAACACCGCCATTTCACAGGGGCTGCTCCGCTGCTGCCCTGCCA 299
Db 146 GGAGCTCTTATGAACACCGCCATTTCACAGGGGCTGCTCCGCTGCTGCCCTGCCA 205
Qy 300 GCGTCACTAGCCGATGAACCCCTTACCTGCCAGGACAGTGAGCTGCTCTGCAATGACTG 359
Db 206 GCGTCACTAGCCGATGAACCCCTTACCTGCCAGGACAGTGAGCTGCTCTGCAATGACTG 265
Qy 360 CTACTGCAGTGGTTCCTGCGAGTGTCCGCTGTTGGGAGACTGTCTATGCTGGGTC 419
Db 266 CTACTGCAGTGGTTCCTGCGAGTGTCCGCTGTTGGGAGACTGTCTATGCTGGGTC 325
Qy 420 CCGGAAGCTGAATATGGAGCCAGACATGGCATGACACTGCTTCTCTGTGCAGTGGCTG 479
Db 326 CCGGAAGCTGAATATGGAGCCAGACATGGCATGACACTGCTTCTCTGTGCAGTGGCTG 385
Qy 480 TGAACAGCCACTGGGCTCCGCTCTTTTGTGCCGACAAAGGCTGCTCACTACTCGCTGCC 539
Db 386 TGAACAGCCACTGGGCTCCGCTCTTTTGTGCCGACAAAGGCTGCTCACTACTCGCTGCC 445
Qy 540 CTGCTATGAGAACAGTTTGTCTCTCGCTCGCGCCGCTGCAGCAAGACGCTGACACAGGG 599
Db 446 CTGCTATGAGAACAGTTTGTCTCTCGCTCGCGCCGCTGCAGCAAGACGCTGACACAGGG 505
Qy 600 TGGAGTACATACCGTATCAGCGGTGGCATCGAGATGTCTGGTCTGTACCGGATGCCA 659
Db 506 TGGAGTACATACCGTATCAGCGATGGCATCGAGATGTCTGGTCTGTACCGGATGCCA 565

Qy 660 GAGCCCTTGCAGGCGAGGAGTTCACCTCCCGGATGAAGATCCCTACTCTGTGGCTG 719
Db 566 GAGCCCTTGCAGGCGAGGAGTTCACCTCCCGGATGAAGATCCCTACTCTGTGGCTG 625
Qy 720 TTTTGGAGAACTCTTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCAATCGTAGACTCGG 779
Db 626 TTTTGGAGAACTCTTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCAATCGTAGACTCGG 685
Qy 780 TGGAGCAAGTATGTCTTTTGAAGACCGACACTGGACCAACAACTGCTTCTCTCTGGC 839
Db 686 TGGAGCAAGTATGTCTTTTGAAGACCGACACTGGACCAACAACTGCTTCTCTCTGGC 745
Qy 840 CCGTCTCTTACCTCCCTGTGGTGGCCAGGGCTTCTTACCGGATGGAGCAAGTGTCTG 899
Db 746 CCGTCTCTTACCTCCCTGTGGTGGCCAGGGCTTCTTACCGGATGGAGCAAGTGTCTG 805
Qy 900 CCAGGGCTGTAGCAGCAGGCGCCCTAAGCAGGGCTCTTGGACCCAGGCTTTCCCATAC 959
Db 806 CCAGGGCTGTAGCAGCAGGCGCCCTAAGCAGGGCTCTTGGACCCAGGCTTTCCCATAC 865
Qy 960 CAGGGGCCAGGACTGTGGCTCTTTTCTAAACCACTCTGGGACTCAGTCTCCCGCCGCC 1019
Db 866 CAGGGGCCAGGACTGTGGCTCTTTTCTAAACCACTCTGGGACTCAGTCTCCCGCCGCC 925
Qy 1020 AAAAAAATGGTCTCTTCTGGCTCCAGGATTTCTCCCACTCCAGCATCCCAAAAC 1079
Db 926 AAAAAAATGGTCTCTTCTGGCTCCAGGATTTCTCCCACTCCAGCATCCCAAAAC 985
Qy 1080 TGGTACTCTCTGACCCAGGCGCCCAATCCTGGGCTTTACAGAGCCTCCATGAGTCAAGC 1139
Db 986 TGGTACTCTCTGACCCAGGCGCCCAATCCTGGGCTTTACAGAGCCTCCATGAGTCAAGC 1045
Qy 1140 CCGCTCCCAACACTGGACTCCAGAAATTCACCTCTCCCTGCGAGTCTGGGTTCCAGAC 1199
Db 1046 CCGCTCCCAACACTGGACTCCAGAAATTCACCTCTCCCTGCGAGTCTGGGTTCCAGAC 1105
Qy 1200 TGAGTCTCTCTCCCAAAATCAGGGCTCTAGACCCGAGCGCTCCAAACCTGGACTCTGGAC 1259
Db 1106 TGAGTCTCTCTCCCAAAATCAGGGCTCTAGACCCGAGCGCTCCAAACCTGGACTCTGGAC 1165
Qy 1260 TTAGGCCCCCTTAAATCTAGACTTCTTTATAGGTTTTCAGGTCTCTATGGGTGCGTGG 1319
Db 1166 TTAGGCCCCCTTAAATCTAGACTTCTTTTATAGGTTTTCAGGTCTCTATGGGTGCGTGG 1225
Qy 1320 GAAGTCTCTGAAAGTGAGCTGTTCTCAGGCTTGAGCTGCCCGACCCATCCCGGGTTG 1379
Db 1226 GAAGTCTCTGAAAGTGAGCTGTTTTCAGGCTTGAGCTGCCCGACCCATCCCGGGTTG 1285
Qy 1380 AGGTGTGGGGGACAGATCAGGAGCCCACTGATAAGGGGCCCTAGGGTACAGGGTGTCT 1439
Db 1286 AGGTGTGGGGGACAGATCAGGAGCCCACTGATAAGGGGCCCTAGGGTACAGGGTGTCT 1345
Qy 1440 GCCCAGAGGTGCGCCACCGAGTGTCTTCTCATTTTATTTTATTTTTCAGTCTCATTTT 1499
Db 1346 GCCCAGAGGTGCGCCACCGAGTGTCTTCTCATTTTATTTTATTTTTCAGTCTCATTTT 1405
Qy 1500 TGGGACAGGGGTGAGATTTGGCTCATCCCTTCCAGATTTCTGCAATAAAGCGGTGTGAG 1559
Db 1406 TGGGACAGGGGTGAGATTTGGCTCATCCCTTCCAGATTTCTGCAATAAAGCGGTGTGAG 1465
Qy 1560 GAAGCAGGGGAAAAA 1580
Db 1466 GAAAAA 1486

RESULT 6
AL603790/c
LOCUS Human DNA sequence from clone Rpl1-100H21 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL603790
VERSION AL603790.41 GI:30519815

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 99687)

REFERENCE

Direct Submission

Submitted (09-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On May 10, 2003 this sequence version replaced gi:30350096.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest, except on the rare

occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information

on the WORMPEP database can be found at

<http://www.sanger.ac.uk/projects/C-elegans/wormpep> This sequence

was generated from part of bacterial clone contigs of human

chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping

Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr1>

RP11-100H21 is from the library RP11-11.1 constructed by the group

of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6.

FEATURES

Source

1. 99687
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-100H21"
/clone_lib="RP11-11.1"
26430 a 24186 c 23985 g 25086 t

BASE COUNT

ORIGIN

Query Match 46.4%; Score 915.2; DB 9; Length 99687;
Best Local Similarity 84.0%; Pred. No. 2.3e-238;
Matches 1153; Conservative 0; Mismatches 3; Indels 216; Gaps 2;
QY 413 CTGGGTCCGGGAAGCTGGAATAGGAGCCAGACATGGCATGAGCAGCTGCTTCTGTGCA 472
DB 70254 CAGGGTCCGGGAAGCTGGAATAGGAGCCAGACATGGCATGAGCAGCTGCTTCTGTGCA 70195
QY 473 GTGGGTGTGAACAGCAGCTGGGCTCCGCTTTTGTGCCCCGACAGGCTGCTCACTACT 532
DB 70194 GTGGGTGTGAACAGCAGCTGGGCTCCGCTTTTGTGCCCCGACAGGCTGCTCACTACT 70135
QY 533 GCGTCCCTGCTATGAGAACAGTTTGTCTCGCTGGCGCGCTGCAGCA----- 583
DB 70134 GCGTCCCTGCTATGAGAACAGTTTGTCTCGCTGGCGCGCTGCAGCAAGGTTGGGG 70075
QY 584 ----- 583

DB 70074 CTGGGCCAGCCAGGTGGGGGTGAGTGAGGCCCTGTTTGTGGGTGGGAATCCCACTCATAC 70015
QY 584 -----AGACGCTGACACAGCGTGGAGTGACATACCGTGATCA 620
DB 70014 CTACTAATGGATGCTGCCCCCAGACAGCGTGGAGTGAGTACATACCGTGATCA 69955
QY 621 GCCGTGGCATCAGAAATGCTGTGTACCCGGATGCCAGAGCCCTGGCAGGCGACGA 680
DB 69954 GCCGTGGCATCAGAAATGCTGTGTACCCGGATGCCAGAGCCCTGGCAGGCGACGA 69895
QY 681 GTTCACTTCCCGGGATGAAGATCCCTACTGTGTGGCTCTTTTGGAGAACTCTTTGACAC 740
DB 69894 GTTCACTTCCCGGGATGAAGATCCCTACTGTGTGGCTCTTTTGGAGAACTCTTTGACAC 69835
QY 741 TAAGTGCAGCAGCTGCAAGCGCCCATCGT----- 770
DB 69834 TAAGTGCAGCAGCTGCAAGCGCCCATCGTAGTGGGACAAAGGTTCAAGAACGAGGTGG 69775
QY 771 ----- 770
DB 69774 GTAGGGTGCAGCAAGGGAGTGGGATGTGTATTGGTCAAGTTGAGGGCCCGGACCATCC 69715
QY 771 -----AGGACTCGGTGGAGCAAGTATGTGT 796
DB 69714 CCACATCTGCATAGCTAACTTCCAGCCTCCCTCCAGGACTCGGTGGAGCAAGTATGTGT 69655
QY 797 CTTTGAAGACCGACACTGGCACCACAACTGCTTCTCTCGCCCGCGCTGCTTACTCTCC 856
DB 69654 CTTTGAAGACCGACACTGGCACCACAACTGCTTCTCTCGCCCGCGCTGCTTACTCTCC 69595
QY 857 TGGTGGGCGAGGCTTCGTACCGGATGGAGACCAAGTGTCTGCGCAGGCTCTAGCCAGG 916
DB 69594 TGGTGGGCGAGGCTTCGTACCGGATGGAGACCAAGTGTCTGCGCAGGCTCTAGCCAGG 69535
QY 917 CAGGCGCTTAAGCCAGGCTCTGTGACCCAGGCTTTCCCATACCAAGGCGCCAGGACTGT 976
DB 69534 CAGGCGCTTAAGCCAGGCTCTGTGACCCAGGCTTTCCCATACCAAGGCGCCAGGACTGT 69475
QY 977 GCTCTCTTTTAAACCACTCTGGSACTCCTCCCGCGCAAAATAAATGGGTCTCC 1036
DB 69474 GCTCTCTTTTAAACCACTCTGGSACTCCTCCCGCGCAAAATAAATGGGTCTCC 69415
QY 1037 TTCTGGGCTCCAGGATGTCTCCCACTCCAGCATCCCAAACTGCTACTCCCTGACCCA 1096
DB 69414 TTCTGGGCTCCAGGATGTCTCCCACTCCAGCATCCCAAACTGCTACTCCCTGACCCA 69355
QY 1097 GCGCCCAATCTGGGCTCTTTACAGAGCTTCCATGAGTCAAGCCCTCCCACTGAGG 1156
DB 69354 GCGCCCAATCTGGGCTCTTTACAGAGCTTCCATGAGTCAAGCCCTCCCACTGAGG 69295
QY 1157 ACTCCAGAAATTCACCTCTCCCTGCTGAGTCTGCGGTTCAGAGTCTGCTTCCCAAAA 1216
DB 69294 ACTCCAGAAATTCACCTCTCCCTGCTGAGTCTGCGGTTCAGAGTCTGCTTCCCAAAA 69235
QY 1217 TCAGGGCTCTAGACCCGAGCCCTCCAACTCTGGACTCTGGGACTTGGGCCCCCTTAAATC 1276
DB 69234 TCAGGGCTCTAGACCCGAGCCCTCCAACTCTGGACTCTGGGACTTGGGCCCCCTTAAATC 69175
QY 1277 TAGACTTCTTTATAGTTTTCAGTCTCTATGGTGGCTGGGAAAGTCTTGAAGTGG 1336
DB 69174 TAGACTTCTTTATAGTTTTCAGTCTCTATGGTGGCTGGGAAAGTCTTGAAGTGG 69115
QY 1337 ACTGTCTTCAGGCTTGAGCTTGCCCACTCCCGGCTGGAGGCTGTGGGGGCGACGA 1396
DB 69114 ACTGTCTTCAGGCTTGAGCTTGCCCACTCCCGGCTGGAGGCTGTGGGGGCGACGA 69055
QY 1397 GATCAGAGCCCACTGATAAGGGGCCCTTAGGGTACAGGGTGTGCCCCAGAGGTCGCCAC 1456
DB 69054 GATCAGAGCCCACTGATAAGGGGCCCTTAGGGTACAGGGTGTGCCCCAGAGGTCGCCAC 68995
QY 1457 CGAGTGTCTTCTCATTTTATTTTCACTCCATTTTCCCATAGATGGCAGAGGGGTGAGA 1516
DB -----

Db 68994 CGAGTGTCTCTCATTTTATTTTTCAGTCCATTTTGGCCCATAGATGGCAGAGGGTGAGA 68935
QY 1517 TTGGCTCATCCCTCCAGATCTGCAATAAAGCGGTGTGAGGAAGCAGGG 1568
|||||
Db 68934 TTGGCTCATCCCTCCAGATCTGCAATAAAGCGGTGTGAGGAAGCAGGG 68883
|||||
RESULT 7
AF134772 1146 bp mRNA linear ROD 27-MAY-1999
LOCUS Mus musculus LIM protein (Fhl3) mRNA, complete cds.
DEFINITION AF134772
ACCESSION AF134772
VERSION AF134772.1 GI:4894846
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1146)
TITLE The LIM proteins FHL1 and FHL3 are expressed differently in
skeletal muscle
JOURNAL Biochem. Biophys. Res. Commun. 255 (2), 245-250 (1999)
MEDLINE 99160848
PubMed 10049693
REFERENCE 2 (bases 1 to 1146)
AUTHORS Morgan, M.J.
TITLE Mouse FHL3 is expressed at high levels in adult skeletal muscle
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1146)
AUTHORS Morgan, M.J.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-1999) Orthodontics, Eastman Dental Institute, 256
Gray's Inn Road, London WC1X 8LD, UK
FEATURES
source 1..1146
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
gene 1..1146
/gene="Fhl3"
CDS 172..1014
/gene="Fhl3"
/note="FHL3"
/codon_start=1
/product="LIM protein"
/protein_id="AAD32623.1"
/db_xref="GI:4894847"
/translation="MSEAFDCAKCNESLYGRKIOTDSGPYCPYCDNTFANTCABCQ
QLTHDSRELTFEDRHFGFCRCRSLAGEPFTQDSELLCNDYCSAFSSQCS
ACGETVMPGSRKLEYGQTHWHEHCFLCSGEOPLGSRSFVPDKGAHYCVPCYENKEAP
RCARCSKTLTQGVTYRDOPWHRECLVCTGCTPLAGQFTSRDEDPYCVACFGEFLFA
PKSSCKRPIVLGGGKYSVFEDRHHHNCFCSCARCSSTLVGGGFPDGDQVLCQCS
QAGP"
misc_feature 190..264
/gene="Fhl3"
/note="Region: HL-domain"
misc_feature 289..447
/gene="Fhl3"
/note="Region: LIM-domain 1"
misc_feature 472..630
/gene="Fhl3"
/note="Region: LIM-domain 2"
misc_feature 655..807
/gene="Fhl3"
/note="Region: LIM-domain 3"
misc_feature 832..996
/gene="Fhl3"
/note="Region: LIM-domain 4"
BASE COUNT 207 a 394 c 313 g 232 t
ORIGIN

Query Match 44.2%; Score 871.6; DB 10; Length 1146;

Best Local Similarity 92.1%; Pred. No. 1.8e-226;
Matches 930; Conservative 0; Mismatches 79; Indels 1; Gaps 1;
QY 13 GTGGGAGAGCTCGCGGAGCTCGCTCGGGCCCCCGCTTGGCGGGCCCGCAGGGTCTCT 72
Db 99 GGCTCCGCGCCACCCACGCGCGCGCGCTCCGCGCGCTTCGAGGGCTCACT 158
|||
QY 73 -CCCTTTGCCACCATGAGCGAGTCAATTGACTGTGCAAAATGCAACGAGTCCCTGTATGG 131
|||
Db 159 GGGCAGCGCCACCATGAGCGAGCATTTGACTGTGCAAAATGCAACGAGTCCCTGTATGG 218
|||
QY 132 AGCAAGTATACATCCAGACAGACAGCGGCCCTACTGTGTCCTGTATGCAATACCTT 191
|||
Db 219 CGCAAAATACATCCAGACAGACAGTGGCCCTACTGCTTCCGCTGTATGCAACACCTT 278
|||
QY 192 TGCCAAACCTGTGTGCTGAGTCCAGCAGCTTATCGGCGATGCTCGAGGAGCTGTCTTA 251
|||
Db 279 CGCAACACCTGTGTGAGTCCAGCAGCTTATCGGCGATGCTCGAGGAGCTGTCTTA 338
|||
QY 252 TGAAGACCGCATTTCCACGAGGCTGCTTCCGCTGCTCGGCTGCCAGCGCTCACTAGC 311
|||
Db 339 TGAAGACCGCATTTCCACGAGGCTGCTTCCGCTGCTCGGCTGCCAGCGCTCACTAGC 398
|||
QY 312 CGATGAACCTTCACTGCGCAGACAGTGTGCTGTGCAATGACTGTCTACTGCACTGC 371
|||
Db 399 CGGTGAACCTTCACTGCGCAGACAGTGTGCTGTGCAATGACTGTCTACTGCACTGC 458
|||
QY 372 GTTTTCTCGCAGTGTCCGCTTGTGGGAGAGTGTATCGCTGGGTCCGCGAGCTGGA 431
|||
Db 459 GTTTTCTCGCAGTGTCCGCTTGTGGGAGAGTGTATCGCTGGGTCCGCGAGCTGGA 518
|||
QY 432 ATATGAGGCGCAGACATGGCATGAGCACTGCTTCTGTGTCAGTGGCTGTGAACAGCCACT 491
|||
Db 519 ATATGAGGCGCAGACATGGCATGAGCACTGCTTCTGTGTCAGTGGCTGTGAACAGCCACT 578
|||
QY 492 GGGCTCCGTTCTTTGTGTCGCCGACAGGCTGTCTACTGTGGTGCCTGTCTATGAGAA 551
|||
Db 579 GGGCTCCGTTCTTTGTGTCGCCGACAGGCTGTCTACTGTGGTGCCTGTCTATGAGAA 638
|||
QY 552 CAAGTTTGTCTCTGCTGCGCTGCGCAGCAGAGCTGTGTCAGAGGCTGTGAGTACATA 611
|||
Db 639 CAAGTTTGTCTCTGCTGCGCTGCGCAGCAGAGCTGTGTCAGAGGCTGTGAGTACATA 698
|||
QY 612 CCGTGATCAGCGTGGCATCGAAGTGTCTGTGTCGTACCGGATGCCAGAGCCGCTTGGC 671
|||
Db 699 CCGTGATCAGCGTGGCATCGAAGTGTCTGTGTCGTACCGGATGCCAGAGCCGCTTGGC 758
|||
QY 672 AGGCGACAGTTCACTCCCGGATGAAGATCCCTACTGTGTGGCTGTCTTTGGAGAACT 731
|||
Db 759 AGGCGACAGTTCACTCCCGGATGAAGATCCCTACTGTGTGGCTGTCTTTGGAGAACT 818
|||
QY 732 CTTTGACCTTAAGTGAGCAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGCAACTA 791
|||
Db 819 CTTTGACCTTAAGTGAGCAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGCAACTA 878
|||
QY 792 TGTGTCTTTGAAGACCGACACTGGCACCACTGCTTCTCTCGCCCGCTGCTCTAC 851
|||
Db 879 TGTGTCTTTGAAGACCGACACTGGCACCACTGCTTCTCTCGCCCGCTGCTCTAC 938
|||
QY 852 CTCCTGTGTGGCGCAGGGCTTCGTACCGGATGGAGACCAAGTGTCTTCCAGGGCTGTAG 911
|||
Db 939 CTCCTGTGTGGCGCAGGGCTTCGTACCGGATGGAGACCAAGTGTCTTCCAGGGCTGCAG 998
|||
QY 912 CAGGCGAGGCGCCCTAAGCCAGGGCTCTGTGACCCAGGCTTCCATATACCAGGCGCCAGG 971
|||
Db 999 CCAAGCAGGCGCCCTGTGAGTGGAGGATCTGTACTCTCTCTTCCCAAAAGAGGTTCAG 1058
|||
QY 972 ACTGTGGCTCTCTTTCTTAAACCACTCTGTGGAGTCACTCTCCCTCCCGCCGCAA 1021
|||
Db 1059 ACTTTGGCGCTCTTTCTTGAACCACTCTGTGGGCGCCAAACCCCTCTCCCTAGAA 1108
|||

RESULT 8

JOURNAL

Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA
 This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal-tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'gg' after last codon and before HindIII site to maintain reading frame.
 Clone distribution: http://bioinfo.clontech.com/orfclones.

FEATURES

source

```

1..843
    location/Qualifiers
      1..843
        /organism="synthetic construct"
        /mol_type="mRNA"
        /db_xref="taxon:32630"
        /clones="GH00305LI.0"
        /clone_lib="BD Creator(TM) CDS Library derived from MGC collection"
        /lab_host="DH5alpha T1 resistant"
        /note="Vector: pDNR-Dual"
      1..>843
        /note="Mutations: 842:Stop->Leu"
        /codon_start=1
        /transl_table=11
        /product="Homo sapiens four and a half LIM domains 3"
        /protein_id="AAP36476.1"
        /db_xref="GI:30584447"
        /translation="MSEFDCAKNESLYGRKYIQTDSGYPVCYPCVNTFANTCAECQ
        OLIGHDSRELFEYDRHFEHFCRCRCORSLADEPFTCDSELICNDYCSAFSSQCS
        ACGETVMPGSRKLEYGGOTWHEHCFILCSGCEPLGSRSPVDPKGAHVCVPCYENKFP
        KRCARSTLTGGVYRDQPHRECLVCTGCTPLAGQFTSRDEDPYCVACGELEFA
        PKCSSKRPITGLGGKRYVSPEDRHWHHNSCFSCARSTSLVGQFVDPDQVLCQCGS
        QAGPL"
  
```

CDS

BASE COUNT
 ORIGIN

163 a 251 c 243 g 186 t

Query Match

Best Local Similarity 42.7%; Score 841; DB 12; Length 843;

Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 85 ATGAGCGAGTCATTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGCAAGTACATC 144
    |||||||
Db 1 ATGAGCGAGTCATTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGCAAGTACATC 60

QY 145 CAGACAGACAGCGCCCTACTGTGTCCTGTATGACAAATACCTTTGCCAACACCTGT 204
    |||||||
Db 61 CAGACAGACAGCGCCCTACTGTGTCCTGTATGACAAATACCTTTGCCAACACCTGT 120

QY 205 GCTGAGTGCCAGCAGCTTATCGGGCATGACTCGAGGAGGCTGTCTATGAGACCGGCAT 264
    |||||||
Db 121 GCTGAGTGCCAGCAGCTTATCGGGCATGACTCGAGGAGGCTGTCTATGAGACCGGCAT 180

QY 265 TTCCACGAGGCGTCTTCCGCTGTGCGGTGCGCGTGCAGGCGCTCAGTAGCCGATGAACCCCTTC 324
    |||||||
Db 181 TTCCACGAGGCGTCTTCCGCTGTGCGGTGCGCGTGCAGGCGCTCAGTAGCCGATGAACCCCTTC 240

QY 325 ACCTGCCAGGACAGTGCCTCTGCAAGACTGCTACTGCACTGGGCTTTTCCTCGCAG 384
    |||||||
Db 241 ACCTGCCAGGACAGTGCCTCTGCAAGACTGCTACTGCACTGGGCTTTTCCTCGCAG 300

QY 385 TGCTCCGCTTGTGGGAGACTGTCATCGTGGGTCCCGGAGCTGGAATATGGAGCCAG 444
    |||||||
Db 301 TGCTCCGCTTGTGGGAGACTGTCATCGTGGGTCCCGGAGCTGGAATATGGAGCCAG 360

QY 445 ACATGGCATGAGCACTGCTTCTGTGAGTGGCTGTGAACAGCCACTGGGCTCCCGTTCT 504
    |||||||
Db 361 ACATGGCATGAGCACTGCTTCTGTGAGTGGCTGTGAACAGCCACTGGGCTCCCGTTCT 420

QY 505 TTTGTGCCGCAAGGCTGCTACTACTGCGTGCCTGCTATGAGACAAGTTTGTCTCT 564
    |||||||
Db 421 TTTGTGCCGCAAGGCTGCTACTACTGCGTGCCTGCTATGAGACAAGTTTGTCTCT 480
  
```

```

QY 565 CGCTGCGCCGCTGCAGCAAGACGCTGACACAGGGTGGAGTGACATACCGTGATCAGCG 624
    |||||||
Db 481 CGCTGCGCCGCTGCAGCAAGACGCTGACACAGGGTGGAGTGACATACCGTGATCAGCG 540

QY 625 TGGCATCAGAAATGCTGGTCTGTACCGGATGCGCAGACGCCCTGGCAGGCGACAGTTC 684
    |||||||
Db 541 TGGCATCAGAAATGCTGGTCTGTACCGGATGCGCAGACGCCCTGGCAGGCGACAGTTC 600

QY 685 ACCTCCGGGATGAAGATCCCTACTGTGTGGCCTGTTTGGAGAACTCTTTGCACCTAAG 744
    |||||||
Db 601 ACCTCCGGGATGAAGATCCCTACTGTGTGGCCTGTTTGGAGAACTCTTTGCACCTAAG 660

QY 745 TGCAGCAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGCAAGTAGTGTCTCTTTGAA 804
    |||||||
Db 661 TGCAGCAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGCAAGTAGTGTCTCTTTGAA 720

QY 805 GACGACACTGGCACCACAACTGCTTCCTGCGCCGCTGCTACTCCTCGTGGG 864
    |||||||
Db 721 GACGACACTGGCACCACAACTGCTTCCTGCGCCGCTGCTACTCCTCGTGGG 780

QY 865 CAGGCTTCTACCGGATGGAGACCAAGTGTCTGCGAGGCTGTAGCCAGGCGGCC 924
    |||||||
Db 781 CAGGCTTCTACCGGATGGAGACCAAGTGTCTGCGAGGCTGTAGCCAGGCGGCC 840

QY 925 T 925
    |
Db 841 T 841

RESULT 10
AY277587
LOCUS Sus scrofa four and a half LIM domains 3 (PHL3) mRNA MAM 12-MAY-2003
DEFINITION Sus scrofa four and a half LIM domains 3 (PHL3) mRNA, complete cds.
ACCESSION AY277587
VERSION AY277587.1 GI:30526304
KEYWORDS
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
  Zuo,B., Xiong,Y.Z. and Deng,C.Y.
  Characterization of the porcine four and a half LIM domains 3
  Unpublished
  2 (bases 1 to 997)
  Zuo,B., Xiong,Y.Z. and Deng,C.Y.
  Direct Submission
  Submitted (12-APR-2003) College of Animal Science and Veterinary
  Medicine, Huazhong Agriculture University, Agricultural Ministry
  Key Laboratory of Swine Genetics and Breeding, Loin Mountain Street
  No.1, Wuhan, Hubei 430070, P. R. China
  Location/Qualifiers
    1..997
      /organism="Sus scrofa"
      /mol_type="mRNA"
      /db_xref="taxon:9823"
      /tissue_type="skeletal muscle"
      /note="breed: Landrace"
    1..997
      /gene="PHL3"
    7..849
      /gene="PHL3"
      /codon_start=1
      /product="four and a half LIM domains 3"
      /protein_id="AAP32084.1"
      /db_xref="GI:30526305"
      /translation="MSEFDCAKNESLYGRKYIQTDSGYPVCYPCVNTFANTCAECQ
      OLIGHDSRELFEYDRHFEHFCRCRCORSLADEPFTCDSELICNDYCSAFSSQCS
      ACGETVMPGSRKLEYGGOTWHEHCFILCSGCEPLGSRSPVDPKGAHVCVPCYENKFP
      KRCARSTLTGGVYRDQPHRECLVCTGCTPLAGQFTSRDEDPYCVACGELEFA
      PKCSSKRPITGLGGKRYVSPEDRHWHHNSCFSCARSTSLVGQFVDPDQVLCQCGS
      QAGP"
  
```

BASE COUNT	181 a	339 c	281 g	196 t	
ORIGIN					
Query Match	41.0%; Score 807.6; DB 4; Length 997;				
Best Local Similarity	89.2%; Pred. No. 5.3e-209;				
Matches	896; Conservative 0; Mismatches 99; Indels 9; Gaps 2;				
QY	79 GCACCATGAGCGAGTCAATTTGACTGTGCAAAATGCAAGAGTCCCTGTATGACGACAG 138				
DB	1 GCACCATGAGCGAGACCTTTGACTGTGCAAAATGCAAGAGTCCCTGTATGACGACAG 60				
QY	139 TACATCCAGACAGACGCGCCCTACTGTGTGCCCTGTATGACAAATFACCTTTGCCAC 198				
DB	61 TACATTACAGACGACACGCGCCCTACTGTGTGCCCTGTATGACAGACCTTTGCCAAC 120				
QY	199 ACCTGTGCTGAGTGCCAGCAGCTTATCGGGCATGACTCGAGGAGCTGTTCTATGAAGAC 258				
DB	121 ACGTGGCTGAGTGCCAGCAGCTTATCGGGCATGACTCAAGGAGCTGTTTACGAGGAT 180				
QY	259 CGCCATTTCCAGAGGGCTGCTTCCGCTGCTGCCGCTGCCAGCGCTCACTAGCCGATGAA 318				
DB	181 CGCCACTTCCAGAGGGCTGTTTCCGCTGCTGCCGCTGCCAGCGCTCCCTCGCCGACGAG 240				
QY	319 CCCTTACCTGCCAGCAGTGAAGTGTCTGCAATGACTGTACTGTGAGTGGGTTTCC 378				
DB	241 CCCTTACCTGCCAGTATGAGCAGCTGCTGTACGACTGTACTGTGACGCGCTTCTCC 300				
QY	379 TCGACTGTCTCGCTTGTGGGAGACTGTCTGCTGCTGCTGCCGAGCTGGAATATGA 438				
DB	301 TCTCAGTGTCTCGCTTGCAGAGACGCTGTATGCTTGGGTCCCGAAGCTGGATACGGA 360				
QY	439 GGCCAGCATGCGATGAGCAGCTGCTTCTGTGTGAGTGGTGTGTAACAGCAGCTGGGCTCC 498				
DB	361 GGCCAGCAGTGGCATGAGCAGCTGCTTCTGTGTGAGCAGGCTGTGAGCAGCGCTGGGCTCC 420				
QY	499 CGTTCTTTTGTGCCACAGAGGTGTCTACTACTGTGCTGCTGCTGCTATGAGAACAGTTT 558				
DB	421 TGTCTCTTTTGTGCCACAGAGGTGTCTACTACTGTGCTGCTGCTGCTATGAGAACAGTTT 480				
QY	559 GTCTCTCGTGTGCCCGCTGCAGACAGCGCTGACACAGGGTGGAGTGACATACCGTGAT 618				
DB	481 GCGCCTCGTGTGCCCGCTGCAGACAGCGCTGACACAGGGTGGAGTGACATACCGTGAT 540				
QY	619 CAGCCTGGATGAGAAATGTCTGTCTGTACCGGATGCCAGAGCGCCCTGGCAGGGCAG 678				
DB	541 CACCCCTGGCATCGAATGCTGTCTGTGACAGAGGTGCCAGACGCGCCCTGGCAGGGCAG 600				
QY	679 CAGTTACCTCCGGGATGAATGCTCTACTGTGTGCGCTGTTTGGAGAACTTTTGA 738				
DB	601 CAGTTACCTCCGGGACGATGATCTTATGTGTGCGCTGTTTGGAGAACTTTTGA 660				
QY	739 CCTAAGTCAGCAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGGCAAGTATGTGTCC 798				
DB	661 OCCAAGTCAGCAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGGCAAGTATGTGTCC 720				
QY	799 TTGGAAGCCGACACTGGCACCACAACTGCTTCTCTGCGCCGCTGTCTACCTCCGTC 858				
DB	721 TTGGAAGCCGACACTGGCACCACAACTGCTTCTCTGCGCCGCTGTCTACCTCCGTC 780				
QY	859 GTGGGCCAGGGCTTCTGTACCGATGGAGACCAAGTGTCTGCCAGGGCTGTAGCCAGGCA 918				
DB	781 GTGGGTCAAGGGTCTGTGCGGATGGAGACCAAGTGTCTGTGCCAGGGTGTAGCCAGGCA 840				
QY	919 GGGCCCTTAAGCAGGGCTCTCTGGACCCAGGCTTTTCCCATACACAGCGCCGACGACTGTGG 978				
DB	841 GGGCCCTTAAGCAGGGCTCTCTGGACCCAGGCTTTTCCCATACACAGGGCTCCAGGACTAAG 900				
QY	979 CTCCTTTTCTAAACACCTCTGGGACTCAGCTCCCGCCGCAAAAAAATGGTGTCTCCPT 1038				
DB	901 CTCCTTTTCTAAACACCTCTGGGACTCAGCTCCCGCCGCAAAAAAATGGTGTCTCCPT 953				
QY	1039 CTGGGCTCCAGGATTT--GTCTCCCACTCCAGCATCCCAACT 1080				
DB	1039 CTGGGCTCCAGGATTT--GTCTCCCACTCCAGCATCCCAACT 1080				
Db	954 CTAGGCTCAGGATTCAGCCTCCCACTCCCACTCCCACTCCCAATCT 997				
RESULT 11					
AF114382					
LOCUS					
DEFINITION	Mus musculus four and half LIM domain protein 3 (Fhl3) mRNA, complete cds.				
ACCESSION	AF114382				
VERSION	AF114382.1				
KEYWORDS	GI:5825392				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 1147)				
TITLE	Chu P.H., Ruiz-Lozano P., Zhou Q., Cai C. and Chen J.				
JOURNAL	Expression patterns of FHL/SLIM family members suggest important functional roles in skeletal muscle and cardiovascular system				
MEDLINE	Mech. Dev. 95 (1-2), 259-265 (2000)				
PUBMED	20368180				
REFERENCE	10906474				
AUTHORS	2 (bases 1 to 1147)				
TITLE	Chen, J.				
JOURNAL	Direct Submission				
FEATURES	Submitted (17-DEC-1998) Medicine, UCSD, 9500 Gilman Drive #5024, San Diego, CA 92093-0613C, USA				
Source	Location/Qualifiers				
gene	1..1147				
CDS	/organism="Mus musculus"				
	/mol_type="mRNA"				
	/db_xref="taxon:10090"				
	1..1147				
	/gene="Fhl3"				
	146..1015				
	/gene="Fhl3"				
	/codon_start=1				
	/product="four and half LIM domain protein 3"				
	/protein_id="AAD53231.1"				
	/db_xref="GI:5825393"				
	/translation="MSEAFDCAKNESLYGRKYIOTDGPYCPVCVPCVNTFANTCAECQ				
	QLIGHDSRELFYEDRHFHEGFCRCORSLADEPTCODSELLNECYCTAFSSQCS				
	ACGETVMSRKLEYGGQVHWHRECLVSCGEQPLASRSFVDPKRAHICVPCYENKFPAP				
	RCARSKLTQGGVYRDQVHWHRECLVSCGEQPLASRSFVDPKRAHICVPCYENKFPAP				
	PKSSCNRPITGGSGGAGAGLGGKIVSFEDRHHWHSFCARCSSTLVGGFVPDG				
	DVLQCGGCSQAGP"				
BASE COUNT	206 a	381 c	319 g	241 t	
ORIGIN					
Query Match	36.1%; Score 711; DB 10; Length 1147;				
Best Local Similarity	81.9%; Pred. No. 1.2e-182;				
Matches	844; Conservative 0; Mismatches 160; Indels 27; Gaps 1;				
QY	18 CAGCCTCGCGGCGAGTCTCGTGGCCCCCGCTTGGCCCCCGCGAGGTTCTCTCCCT 77				
DB	79 CGGCTCGCGGCGAGTCTCGTGGCCCCCGCTTGGCCCCCGCGAGGTTCTCTCCCT 138				
QY	78 TGCCACCATGAGCAGTCAATTTGACTGTGCAAAATGCAACAGTCCCTGTATGAGGACAA 137				
DB	139 CGGCACCATGAGCAGTCAATTTGACTGTGCAAAATGCAACAGTCCCTGTATGAGGACAA 198				
QY	138 GTACATCCAGACAGCAGCGCCCTACTGTGTGCCCTGTATGACAAATACCTTTGCCAA 197				
DB	199 ATACATCCAGACAGCAGTGGCCCCCTACTGTGCTGCTATGACACACCTTCGCCAA 258				
QY	198 CACCTGTGTGATGCCAGCAGCTTATCGGGCATGACTCGAGGAGCTTTCTATGAAGA 257				
DB	259 CACCTGTGTGATGCCAGCAGCTTATCGGGCATGACTCGAGGAGCTTTCTATGAAGA 318				
QY	258 CGGCATTTCCAGAGGGTCTCGTGGCTGCTGCGCTGCCGCTGCCAGGCTACTAGCGATCA 317				
DB	319 CCGCCACTTCCAGAGGGTCTCGTGGCTGCTGCCGATGCCAGGCTCTCTTGGCGATGA 378				

DEFINITION Mus musculus chromosome 6 clone RP23-4015 strain C57BL6/J, complete sequence.

ACCESSION AC027285

VERSION AC027285.4 GI:20486398

KEYWORDS HTG

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Avele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-O., Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Masello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A., Shevchenko, Y., Snyder, B., Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Zhang, L.-H. and Green, E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 241886)

Green, E.D.

Direct Submission

Submitted (29-MAR-2000) NIH Intramural Sequencing Center, 8717

Grovmont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 241886)

Green, E.D.

Direct Submission

Submitted (09-AUG-2001) NIH Intramural Sequencing Center, 8717

Grovmont Circle, Gaithersburg, MD 20877, USA

4 (bases 1 to 241886)

Green, E.D.

Direct Submission

Submitted (07-MAY-2002) NIH Intramural Sequencing Center, 8717

Grovmont Circle, Gaithersburg, MD 20877, USA

On May 7, 2002 this sequence version replaced gi:15145103.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc_mouse@nhgri.nih.gov

----- Project information

Center project name: xg

Center clone name: 040105

This sequence was finished as follows unless otherwise noted:

all regions were double-stranded, sequenced with an

alternate chemistry, or covered by high quality data

(i.e., phred quality >= 30); an attempt was made to resolve

all sequencing problems, such as compressions and repeats;

all regions were covered by at least one plasmid subclone

or more than one M13 subclone; and the assembly was confirmed

by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of

this clone unless otherwise noted. If there are overlapping

clones, the overlaps are noted in the beginning and end of

the Features section.

Location/Qualifiers

1..241886

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL6/J"

/db_xref="taxon:10090"

/chromosome="6"

/clone="RP23-4015"

/clone_lib="RPCI mouse BAC library 23"

181995..182091

/note="unresolved tandem repeat; approximately 128 bp.

missing"

205400..205425

/note="unresolved tandem repeat; approximately 184 bp

missing"

BASE COUNT 67655 a 48051 c 50506 g 75674 t

Query Match 20.4%; Score 402.2; DB 10; Length 241886;
Best Local Similarity 64.7%; Pred. No. 2.5e-98;
Matches 829; Conservative 0; Mismatches 263; Indels 189; Gaps 8;

QY 428 TGAATATGAGGCGGAGAGCATGCGATGAGCACTGCTCTCTGTCAGTGTCTGTGCAAGCAGC 487

98821 TGGAGTATGAGGCGGAGAGCGTGCATGACATGCTTTCTGTGCAAGTGTCTGTGAGCAGC 98762

QY 488 CACTGGCTCCCGTCTCTTTTGTGCCGAGCAAGGTGCTCCTACTGCGTGCCTGCTGTATG 547

98761 CGCTGGCTCCCGTCTCTTTTGTGCCGAGCAAGGTGCTCCTACTGCGTGCCTGCTGTATG 98702

QY 548 AGAACAAGTTCTCTCTGCTGCGCGCTGAGCAAGAGCGTGCACAGGTGGAGTGA 607

98701 AGAACAAGTTCTCTCTGCTGCGCGCTGAGCAAGAGCGTGCACAGGTGGAGTGA 98642

QY 608 CATACCGTATCAGCGCGTGCATCGAATGCTGTGTACCGGATGCCAGAGCGCC 667

98641 CATATCGTATCAGCGCGTGCATCGAATGCTGTGTACCGGATGCCAGAGCGCC 98582

QY 668 TGGAGGCGAGCAGTTCACTCCCGGATGAGAGTCCCTACTGCTGTGGCTGTGTTGGAG 727

98581 TTGCAGAGCAGCAGTTCACTCCCGGATGAGAGTCCCTACTGCTGTGGCTGTGTTGGAG 98522

QY 728 AACTCTTTCACCTTAAGTGCAGCAGCTCAAGCGCCCATC----- 768

98521 AACTCTTTCACCTTAAGTGCAGCAGCTCAAGCGCCCATC----- 98462

QY 769 ----- 768

98461 AGGGCTCAGTAAAGTGGGATGTAACACGGGTGAGGTAAAGGGGAGCAGCTGGCC 98402

QY 769 -----GTAGGACTCGGTGGAGGCAAGTATGTGTCC 798

98401 AGACCTGCAGAGCTAACCTCCAGCTTCTCCAGGACTCGGTGGAGCAAGTATGTGTCC 98342

QY 799 TTTGAAGCCGACACTGGCACCACCAACTGCTTCTCTGCGCCGCTGCTTACCTCCCTG 858

98341 TTTGAAGCCGACACTGGCACCACCAACTGCTTCTCTGCGCCGCTGCTTACCTCCCTG 98282

QY 859 GTGGCCAGGGCTCTGCTACCGGATGGAGACCAAGTGTCTGCCAGGGTGTAGCCAGGCA 918

98281 GTGGCCAGGGCTCTGCTACCGGATGGAGACCAAGTGTCTGCCAGGGTGTAGCCAGGCA 98222

QY 919 GGGCCCTAAGCCAGGGCTCTGAGCCAGGGCTTTCCCATACACGGGCGCCAGGACTGTGG 978

98221 GGGCCCTAAGCCAGGGCTCTGAGCCAGGGCTTTCCCATACACGGGCGCCAGGACTGTGG 98162

QY 979 CTCCTTTTAAACCACCTCTGGGACTCAGCTCCCGCCGCAAAATGGTCTCCTT 1038

98161 CGCTTTTAAACCACCTCTGGGACTCAGCTCCCGCCGCAAAATGGTCTCCTT 98102

QY 1039 CTGGGCTCCA-----GGATTGCTCCGACCTCCAGCATCCCAACTGG 1082

98101 CGCTTTTAAACCACCTCTGGGACTCAGGATTCAGTACCACTTTTAAACACCCAGGCTGG 98042

QY 1083 TACTCCCTGAGCCAGGGCGCCCAATCCTGGGCTTTTACAGAGCCTCCATGAGTCAAGCC 1142

98041 TACTCCCTGAGCCAGGGCGCCCAATCCTGGGCTTTTACAGAGCCTCCATGAGTCAAGCC 97983

QY 1143 CTCCTTTTAAACCACCTCTGGGACTCAGCTCCCGCCGCAAAATGGTCTCCTT 1202

97982 TTTCTTAAACCACCTCTGGGACTCAGCTCCCGCCGCAAAATGGTCTCCTT 97937

QY 1203 GTCCTTCTCCCAATCAGGGCTCTAGAGCCGAGCGCTCCAAACCTCGGACTCTGGACTTA 1262

97936 -----ACACCCAGCGCTCCAAACCTCGGACTCTGGACTTA 97901

QY 1263 GGGCCCTTAAATCTAGACTTCTCTTTTATAGGTTCAGGCTCTCCTATGGTGCTGGG 1320

Db 97900 GGCTCCTTAAGTCAAACTTCTAGGCTAGATTTTTTTAGTTTCTGCGAGTGCCTAAG 97841

QY 1321 AAGTCCCTTGAAGTGGACTCTTCTCAGGCTT-----GACCTGCCCCACCCCATCCCC 1372

Db 97840 AAGTCTTCAAAATAGATTGTATACACCGGTTTGACCTCACCCTCCACTGCTCTCTC 97781

QY 1373 GCGGTTGAGGCTGTGGGGGAGGATGATGAGGAGCCACTGATAGGGGGCCCTAGGGTACA 1432

Db 97780 TGAGTCTGCGCCATATGATAGCAGATTAGGGCTCTGCTAGAGGGTTGTAAGGTACA 97721

QY 1433 GGGTGTGCGCCAGGAGTCCCGAGGAGTCTCTTCATTTTATTCAGTCCCATTTTC 1492

Db 97720 GGGTGTGCTCCAGCCTCTGAGTAC-----CATGTTATTTTCAGTCCCATTTTC 97673

QY 1493 CCATAGATGGCAGAGGGGTGAGATTGCTCATCCCTCCAGATTCTCAATAAAGCG 1552

Db 97672 CCTGCAACAGAGCAG---GAAGGTGTTTACCCCTGTTTCCAGAGATTGTAATAAGCA 97616

QY 1553 GTGTGAGGAAGCAGGGAAAAA 1573

Db 97615 GTGTGAGGGGGGAAAAA 97595

RESULT 14

AB008571

LOCUS

DEFINITION Rattus norvegicus mRNA for DRAL, complete cds.

ACCESSION AB008571

VERSION AB008571.1 GI:2605503

KEYWORDS DRAL.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

REFERENCE 1 (sites) Tanahashi, H. and Tabira, T. Alzheimer's disease-associated presenilin 2 interacts with DRAL, an LIM-domain protein Hum. Mol. Genet. 9 (15), 2281-2289 (2000)

JOURNAL MEDLINE 20458893

REFERENCE 2 (bases 1 to 840) Tanahashi, H. Direct Submission Submitted (01-NOV-1997) Hiroshi Tanahashi, National Institute of Neuroscience, NCNP, Division of Demyelinating Disease and Aging; 4-1-1 Ogawahigashi, Kodaira, Tokyo 187-8502, Japan (E-mail: tanahashencn@ncnp.go.jp, Tel: 81-42-341-1717, Fax: 81-42-346-1747)

FEATURES

source Location/Qualifiers

1..840

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Wistar"

/db_xref="taxon:10116"

/tissue_type="brain"

1..840

/codon_start=1

/product="DRAL"

/protein_id="BAA23357.1"

/db_xref="GI:2605504"

/translations="MVERFDCHNCSNLSYKVKYLKRNPHCVACPEELVANTCEEG TPICGCKDLISYKDRWHQECFHCSCGSSLDVKPRAKEQLCLTDCYSNEYSSKQ ECKTIPGTRKMEYKSWHETCFQCQRCQOPKYSFIPKFNQNPFCVPEKQYAL QCVCKPIITGGVTRNDQWHRCECFVCTACKKQLSQRFETARDEFPYCLTFCFDLIA KKCACTNPISGLGGTKYISFEERQHNDCFNCKKCSLSLVGRFLTERDDIILCPDGG KDI"

117..279

/note="encodes LIM-domain 1"

301..459

/note="encodes LIM-domain 2"

484..636

misc_feature /note="encodes LIM-domain 3" 661..825

BASE COUNT 202 a 227 c 234 g 177 t

ORIGIN

Query Match 17.5%; Score 344; DB 10; Length 840;

Best Local Similarity 63.3%; Pred. No. 1.9e-82;

Matches 527; Conservative 0; Mismatches 305; Indels 0; Gaps 0;

QY 85 ATGACGAGTCATTTGACTGTGCAAAATGCAACAGAGTCCCTGTATGAGCAGCAAGTACATC 144

Db 1 ATGACTGAAGCCTTTGACTGTCACCACTGCAATCAGTCTCTCTTATGGCAAGTACATTT 60

QY 145 CAGACACAGCGCCCTACTGTGCTGCTGCTATGACATATCTTTTGGCAACACCTGT 204

Db 61 CTAAGGAGGAGAACCCCGCACTGTGTGCTGCTTTCGAGGAGCTCTATGCCCAACACCTGT 120

QY 205 GCTGAGTGCCAGCAGCTTATCGGGCATGACTCGAGGAGGCTGTCTTATGAAGACCGCAT 264

Db 121 GAGGAGTGTGAACACCACTCGGCTGTGACTGCAAGGACTTGTCTTACAAGGATCGCCAC 180

QY 265 TTCCACAGGAGGCTCTCCGCTGCTGCGCTGCGCAGCGCTCAGTACAGCGATGACACCTTC 324

Db 181 TGGCATGAAGGCTGCTTCCACTGCTCCGGTGGGAGCTGCTGCTGGTGGCAAGACCCCTTT 240

QY 325 ACCTGCCAGCAGTGTGAGTGTCTGCAATGACTGCTACTGCACTGCTGCTTTCCTCGCAG 384

Db 241 GCGGCCAAGGAGGAGCAGCTGCTGTGCAACGACTGCTATTCACAGGACTCTCTGCTCCAAG 300

QY 385 TGCTCCGCTTTGTGGGAGACTGTCTGCTGGGTCCGGAGAGCTGGAATATGAGAGCGCAG 444

Db 301 TGCAGGAGTGTGAAGAGCACTCATGCCAGTACCCAGAGATGGAATACAAGGCGCAGC 360

QY 445 ACATGGCATGAGCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504

Db 361 AGCTGGCAGCAGAGCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 505 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564

Db 421 TTCTATCTTGAAGAGATCAGAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

QY 565 CGCT 624

Db 481 CAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

QY 625 TGGCATGCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684

Db 541 TGGCAGAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 685 ACCTGCCGCGATGAAGATCCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744

Db 601 ACAGCCCGGATGAGTTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 745 TGCAGCAGCTGCAAGCGCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804

Db 661 TGTGCTGGTGCACCAACCCCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

QY 805 GACCGACACTGGCAGCAGCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 864

Db 721 GAACGCCAGTGGCACAACGACTGCTTAACTTGAAGAGTGTCTCTCTGCTGCTGCTGCTGCT 780

QY 865 CAGGCTTCGTACCGGATGAGACCAAGTGTCTGCCAGGCTCTGTAGCAGG 916

Db 781 CGAGGCTTCTCAGAGAGAGACACATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 832

RESULT 15

AF114381

LOCUS

DEFINITION Mus musculus four and half LIM domain protein 2 (Phl2) mRNA, complete cds.

ACCESSION AF114381

```

VERSION AF114381.1 GI:5825390
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 876)
AUTHORS Chu, P.H., Ruiz-Lozano, P., Zhou, Q., Cai, C. and Chen, J.
TITLE Expression patterns of FHL/SLIM family members suggest important
functional roles in skeletal muscle and cardiovascular system
JOURNAL Mech. Dev. 95 (1-2), 259-265 (2000)
MEDLINE 20368180
PUBMED 10906474
REFERENCE 2 (bases 1 to 876)
AUTHORS Chen, J.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1998) Medicine, UCSD, 9500 Gilman Drive #5024,
San Diego, CA 92093-0613C, USA
FEATURES
source
1..876
Location/Qualifiers
1..876
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..876
/gene="Fhl2"
12..851
/gene="Fhl2"
/codon_start=1
/product="four and half LIM domain protein 2"
/protein_id="A053230.1"
/db_xref="GI:5825391"
/translation="MTERFDCHEHNSLYGKRYILKEENPHCVAFEEELYANTCEEG
TPIGDCDLVSKDRHGGCFHSGRSSLDVDPFAAKEQLLICDYSNEYSKCO
ECKTITMPTKMEYKSGSWHETCTCORCOPICTKSFIPKENONFCVPCYKQYAL
QCVCKKPTITGGTYIREQPHKFCFVCTACKKQLSGORFTARBEFPYCLICFCFLYA
KRCACITPISLGLGRTKYSIFERQWHDNCFNCKCSLSVLGRFLTERDDILCPDCG
KDI"
BASE COUNT 219 a 232 c 242 g 183 t
ORIGIN
Query Match 16.9%; Score 333.4; DB 10; Length 876;
Best Local Similarity 62.3%; Pred. No. 1.5e-79;
Matches 523; Conservative 0; Mismatches 316; Indels 0; Gaps 0;
78 TGCCACCATGAGCGAGTCATTGACTGTGCAGAAATGCACAGAGTCCCTGTATGGACGCAA 137
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
5 TGTCAGCATGACTGAACGCTTTGACTGCCACACTGCATGAATCTCTGTACGGCAAGAA 64
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 GTACATCCAGACAGACAGCGGCCCTACTGTGTGCCCTGTGTATGACATACCTTTGCCAA 197
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
65 GTACATTTAAAGGAGGAGAACCCACACTGTGTGGCTGTCTTGTAGAGAACTATATGCCAA 124
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
198 CACCTGTGCTGAGTCCAGCAGCTTATCGGCATGACTCGAGGAGCTGTCTATGAAGA 257
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
125 TACCTGTGAGGAGTGTGAACACCATCGGTGTGACTGCAAGGACTTGTCTCTACAAGGA 184
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
258 CGCCATTTCCAGGAGCTGCTCCGCTGTGCGCTGCCAGCGCTCACTAGCCGATGA 317
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
185 TCGGCACTGGCATGAAGCTGCTTCCACTGTCCCGTGGGGAGCTCGCTGGTGACAA 244
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
318 ACCCTTACCTGCCAGGACAGTGAAGTCTGTGTGAATGACTGCTACTGAGTGCCTTTTC 377
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
245 GCCCTTTGGCCCAAGAGGAGCAGCTGCTGTGCACCGACTGCTATTCCACAGGACTC 304
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
378 CTCGAGTGTCCGCTTCTGGGAGACTGTATGCTGGTCCCGGAGCTGGAATATGG 437
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
305 GTCCAAGTGCAGGAGTGTAGAAGACCATCGGATACCCCGCAAGATGGAATACAA 364
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
438 AGGCCAGACATGGCATGAGCACTGCTTCTGTGAGTGGCTGTGAACCCACTGGGCTC 497
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
365 GGGCAGCAGCTGGCAGGAGACTGCTTACCTGTGAGGCTGCCAGAGCCCATTTGGAAC 424
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
498 CCGTTCCTTTGTGCCCCCAAGAGGAGTGTACTGTGCTGCTGCTGTATGAGAACAAAGTT 557
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: August 27, 2003, 13:24:08
Job time : 4841 secs

```

425 CAAGAGCTTCATACCTTAAGGAGAATCAGAACTTCTGCGTCCCTGCTATGAGAAGCAGTA 484
558 TGCTCCTCGCTGCGCCCGCTGCAGCAAGACGCTGCACAGGCTGGAGTGACATACCGTGA 617
485 TCCCTCGCAGTGGTGGCAGTGCAAAAAGCCTATACCAAGAGGAGTGTACTTACCGGGA 544
618 TCAGCCGTGGCATCGAGAATGTCTGTGTACCGGATGCCAGAGCCCGCTGCCAGGCA 677
545 GCAGCCCTGGCAGCAAGGATGCTTGTGTGCACAGCCCTGCAAGAAGCAGCTATCTGGGA 604
678 GCAGTTACCTCCCGGATGAAGATCCCTACTGTGTGGCCTGTTTGGAGAACTCTTTGC 737
605 ACGCTTCACAGCAGCGGATGAGTTTCCATACTGCTGCTGCTGCTGCTGCTGCTGCTG 664
738 ACCTAAGTGCAGCACTGCAAGCGCCCATCTGATAGACTCGTGGAGGCAACTATGTGTC 797
665 TAAGAAATGTGCTGGTGCACCAACCCCATTAGTGGTGTGGGTGGCACAAGATACATC 724
798 CTTTGAAGACCGACACTGCGACCACTGCTTCTCTGCGCCCGCTGCTTACCTCCCT 857
725 TTTGAGGAACGCCAATGGCACAAGACTGCTTTAACTGTAGAAGTGTCTCTCTCT 784
858 GGTGGCCAGGCTTTCGTACCGGATGGAGACCAAGTGTCTGCCAGGGGTGTAGCCAGG 916
785 GGTGGGAGGAGGCTTCTCTACAGAGAGAGATGACATCCTCTGCTGCTGCTGCTGCTG 843

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 10:37:18 ; Search time 394 Seconds
(without alignments)
13504.042 Million cell updates/sec

Title: US-09-786-135-2
Perfect score: 1971
Sequence: 1 ggtcgtcgtcagtcgag.....ctcttgacgtggggatgaa 1971

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq.13Jun03.*			
1:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*		
2:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*		
3:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*		
4:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*		
5:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*		
6:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*		
7:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*		
8:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*		
9:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*		
10:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*		
11:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*		
12:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*		
13:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*		
14:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*		
15:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*		
16:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*		
17:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*		
18:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*		
19:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*		
20:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*		
21:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*		
22:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*		
23:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*		
24:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*		
25:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1971	100.0	1971	21	AAZ99798
2	332.4	16.9	1585	20	AAZ41281
3	296.4	15.0	318	22	AAH81611
4	262.4	13.3	1158	24	ABQ99467
5	255.4	13.0	1673	20	AAZ42134
6	255.4	13.0	2340	24	ABK34596
7	252.2	12.8	2254	24	ABL62451
8	252.2	12.8	2254	24	ABL66934

9	244.4	12.4	2310	22	AAH23101
10	240	12.2	2254	25	ACA56524
11	232.6	11.8	830	22	AAO1353
12	232.6	11.8	830	22	AAH33146
13	232.6	11.8	830	23	ABL98806
14	225	11.4	2042	24	ABT10167
15	225	11.4	2154	25	ABX63260
16	196.4	10.0	239	22	AAH82192
17	176.8	9.0	883	23	ABL15937
18	165	8.4	167	16	AAH21976
19	140.4	7.1	1284	22	AAF25354
20	117.6	6.0	469	21	AAO1535
21	115.8	5.9	723	24	AAH62000
22	107.4	5.4	1311	23	AAH76147
23	92.4	4.7	707	23	ABL18295
24	88.6	4.5	550	21	AAH93589
25	81.2	4.1	1593	23	ABL20585
26	81.2	4.1	1793	23	ABL19665
27	81.2	4.1	12706	23	ABL19664
28	81.2	4.1	14899	23	ABL20584
29	75.6	3.8	1486	24	AAH24709
30	75.4	3.8	1456	21	AAH87741
31	75.4	3.8	1488	20	AAH24321
32	75.4	3.8	1488	21	AAH87714
33	75.4	3.8	1575	21	AAH87742
34	75.4	3.8	1597	17	AAH12908
35	75.4	3.8	1620	21	AAH87737
36	75.4	3.8	1644	20	AAH24322
37	75.4	3.8	1644	21	AAH87715
38	75.4	3.8	1665	20	AAH24328
39	75.4	3.8	1665	21	AAH87738
40	75.4	3.8	1689	20	AAH24327
41	75.4	3.8	1689	21	AAH87729
42	75.4	3.8	1725	24	ABK83879
43	75.4	3.8	1835	21	AAH18013
44	75.2	3.8	7302	23	ABL15936
45	75	3.8	802	24	ABS51464

ALIGNMENTS

RESULT 1
AAZ99798
ID AAZ99798 standard; cDNA; 1971 BP.

XX AAZ99798;

AC AAZ99798;

XX 12-JUL-2000 (first entry)

DE cDNA encoding a human LIM domain protein homologue.

XX Human; LIM domain protein homologue; LIM domain; L protein; I protein;

KW M protein; Incyte clone 1925017; cancer; reproductive disorder;

KW autoimmune disorder; ss.

XX Homo sapiens.

OS Key

XX Location/Qualifiers

FT CDS

FT 85..927

FT /*tag= a

FT /product= "LIM domain protein homologue"

XX WO200014231-A1.

XX PD 16-MAR-2000.

XX PF 01-SEP-1999; 99WO-US20239.

XX PR 03-SEP-1998; 98US-0155260.

XX PA (INCY-) INCYTE PHARM INC.

PI Tang YT, Guegler KJ, Corley NC, Patterson C;
XX WPI: 2000-256982/22.
DR P-PSDB; ANY84378.
XX
PT A purified polypeptide designated LDPH (LIM domain protein homolog),
PT useful for treating, preventing, and diagnosing various cancers,
PT reproductive disorders and autoimmune disorders -
XX
PS Claim 9; Page 53-54; 59pp; English.
XX
CC The present sequence encodes a human LIM domain protein homologue.
CC The LIM domain is so named because it was first described in three
CC proteins from *Drosophila melanogaster*, designated L, I, and M. The LIM
CC domain is rich in cysteine. Nucleic acids encoding the LIM protein
CC homologue were first identified in incyte clone 1925017 from a breast
CC tissue cDNA library. The LIM polynucleotide and polypeptide are
CC useful for treating, preventing, and diagnosing various cancers,
CC reproductive disorders and autoimmune disorders.
XX
SQ Sequence 1971 BP: 399 A; 618 C; 528 G; 426 T; 0 other;

Query Match 100.0%; Score 1971; DB 21; Length 1971;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1971; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCGTCGTGAGTCGGCAGCTCGCGGACAGCTCGCTCGGCCCGCCCTGGCCCGGCC 60
Db 1 GGTCCGTCGTGAGTCGGCAGCTCGCGGACAGCTCGCTCGGCCCGCCCTGGCCCGGCC 60

Qy 61 GCAGGGTTCCTCCCTTGGCCACCATGAGCGAGTCATTGACTGTGCAAAATGCAACGAG 120
Db 61 GCAGGGTTCCTCCCTTGGCCACCATGAGCGAGTCATTGACTGTGCAAAATGCAACGAG 120

Qy 121 TCCCTGTATGACCAAGTACATCCAGACAGACAGCGGCCCTACTGTGCTGCTAT 180
Db 121 TCCCTGTATGACCAAGTACATCCAGACAGACAGCGGCCCTACTGTGCTGCTAT 180

Qy 181 GACAAATACCTTTGCAACACCTGTGCTGAGTGCCAGACAGCTTATCGGGCATGACTCGAGG 240
Db 181 GACAAATACCTTTGCAACACCTGTGCTGAGTGCCAGACAGCTTATCGGGCATGACTCGAGG 240

Qy 241 GAGCTGTTCTATGAAGACCGCATTTCCACGAGGCTGCTTCCGCTGCTGCGCTGCGC 300
Db 241 GAGCTGTTCTATGAAGACCGCATTTCCACGAGGCTGCTTCCGCTGCTGCGCTGCGC 300

Qy 301 CGCTCACTAGCGATGAACCTTCCACCTGCCAGACAGAGTGAGTGTCTGCAATGACTGC 360
Db 301 CGCTCACTAGCGATGAACCTTCCACCTGCCAGACAGAGTGAGTGTCTGCAATGACTGC 360

Qy 361 TACTGCACTGCGTTTCTCGAGTGTCTCGCTTGTGGGAGACTGTCATGCTGGTGC 420
Db 361 TACTGCACTGCGTTTCTCGAGTGTCTCGCTTGTGGGAGACTGTCATGCTGGTGC 420

Qy 421 CGGAAGCTGGAATATGAGGCGACATGGATGAGCACTGCTTCTGTGAGTGGCTGT 480
Db 421 CGGAAGCTGGAATATGAGGCGACATGGATGAGCACTGCTTCTGTGAGTGGCTGT 480

Qy 481 GAACAGCACTGGGCTCCCGTTCCTTGTGCGGACAAAGGTGCTCACTACTGCTGCC 540
Db 481 GAACAGCACTGGGCTCCCGTTCCTTGTGCGGACAAAGGTGCTCACTACTGCTGCC 540

Qy 541 TGCTATGAGAACAAATTTGCTTCCTCGCTGCGGCCGCTGCAGAACAGCGCTGACACAGGT 600
Db 541 TGCTATGAGAACAAATTTGCTTCCTCGCTGCGGCCGCTGCAGAACAGCGCTGACACAGGT 600

Qy 601 GGAGTGACATACCTGTATCAGCCGTGGCATGCGAATGTCTGTCTGTACCGGATGCCAG 660
Db 601 GGAGTGACATACCTGTATCAGCCGTGGCATGCGAATGTCTGTCTGTACCGGATGCCAG 660

Qy 661 ACGCCCTTGGCAGGCGAGCTTCCACCTCCCGGATGAAGATCCCTACTGTGTCCTGT 720
Db 661 ACGCCCTTGGCAGGCGAGCTTCCACCTCCCGGATGAAGATCCCTACTGTGTCCTGT 720

Qy 721 TTTGGAGAACTCTTTTGACACCTAAGTCAGCAGCTGCAAGCGCCCATCGTAGGACTCGGT 780
Db 721 TTTGGAGAACTCTTTTGACACCTAAGTCAGCAGCTGCAAGCGCCCATCGTAGGACTCGGT 780

Qy 781 GGAGCGAAGTATGTCTCTTTGAAGACCGACACTGGCACCACAACTGCTTCTCTCGGCC 840
Db 781 GGAGCGAAGTATGTCTCTTTGAAGACCGACACTGGCACCACAACTGCTTCTCTCGGCC 840

Qy 841 CGCTGCTCTACCTCCCTGCTGGGCCAGGGCTTCTACGGATGGAGACCAAGTCTCTGC 900
Db 841 CGCTGCTCTACCTCCCTGCTGGGCCAGGGCTTCTACGGATGGAGACCAAGTCTCTGC 900

Qy 901 CAGGGCTGTAGCAGGAGGCGCCCTAAGCCAGAGCTCTCTGACCCAGCTTTTCCCATACC 960
Db 901 CAGGGCTGTAGCAGGAGGCGCCCTAAGCCAGAGCTCTCTGACCCAGCTTTTCCCATACC 960

Qy 961 ACGGGCCAGGACTGTGGCTCCTTTTCTAAACCACTCTGGGACTCAGCTCCCCCGCCA 1020
Db 961 ACGGGCCAGGACTGTGGCTCCTTTTCTAAACCACTCTGGGACTCAGCTCCCCCGCCA 1020

Qy 1021 AAAAAATGGTCTCTTCTGGGCTCCAGGATTTGCTCCCCACATCCAGCATCCCCAACT 1080
Db 1021 AAAAAATGGTCTCTTCTGGGCTCCAGGATTTGCTCCCCACATCCAGCATCCCCAACT 1080

Qy 1081 GGTACTCCCTGACCCAGGCGCCCAATCTTGGGCTCTTACAGAGCTCCATGAGTCAAGCC 1140
Db 1081 GGTACTCCCTGACCCAGGCGCCCAATCTTGGGCTCTTACAGAGCTCCATGAGTCAAGCC 1140

Qy 1141 CCCTCCCGACACTGGACTCCAGAAATTCACCTCTCCCTCGCATCTGGGTTCACAGACT 1200
Db 1141 CCCTCCCGACACTGGACTCCAGAAATTCACCTCTCCCTCGCATCTGGGTTCACAGACT 1200

Qy 1201 GAGTCTCTCCCAAAATCAGGGCTCTAGACCCGAGCCCTCCAAACCTGGGACTCTGGGACT 1260
Db 1201 GAGTCTCTCCCAAAATCAGGGCTCTAGACCCGAGCCCTCCAAACCTGGGACTCTGGGACT 1260

Qy 1261 TAGSCCCCTTAAATCTAGACTTCTCTTTATAGTCTTTCAGGCTCTCTATGGGTGCTGGG 1320
Db 1261 TAGSCCCCTTAAATCTAGACTTCTCTTTATAGTCTTTCAGGCTCTCTATGGGTGCTGGG 1320

Qy 1321 AAGTCTTTGAAAGTGGACTGTCTCAGGCTTGACCTGCCCCACCCCATCCCCGGGTGTA 1380
Db 1321 AAGTCTTTGAAAGTGGACTGTCTCAGGCTTGACCTGCCCCACCCCATCCCCGGGTGTA 1380

Qy 1381 GGCTGTGGGCGACAGATCAGGAGCCACATGATAAGGGGCGCTAGGCTACAGGTGCTG 1440
Db 1381 GGCTGTGGGCGACAGATCAGGAGCCACATGATAAGGGGCGCTAGGCTACAGGTGCTG 1440

Qy 1441 CCCAGCAGGTGCGCACCGAGTGTCTCTCAATTTTATTCAGTCCATTTTGGCCCATAGAT 1500
Db 1441 CCCAGCAGGTGCGCACCGAGTGTCTCTCAATTTTATTCAGTCCATTTTGGCCCATAGAT 1500

Qy 1501 GGGCAGAGGGTGGATTTGGCTCATCCCCCTTCCAGATTTCTGCAATTAAGGGGTGTGAGG 1560
Db 1501 GGGCAGAGGGTGGATTTGGCTCATCCCCCTTCCAGATTTCTGCAATTAAGGGGTGTGAGG 1560

Qy 1561 AAGCAGGAAAAAAGGGGCGCGCGCTATTGGACTCGCGACCGCGGAAATAA 1620
Db 1561 AAGCAGGAAAAAAGGGGCGCGCGCTATTGGACTCGCGACCGCGGAAATAA 1620

Qy 1621 TTTCCGCGCGGTCCCTTGGGAGGAAACAGTGTTCCTTAAAGTGGCTCGTTTAAAAACC 1680
Db 1621 TTTCCGCGCGGTCCCTTGGGAGGAAACAGTGTTCCTTAAAGTGGCTCGTTTAAAAACC 1680

Qy 1681 TGGGGGAAAAACAGGGGTAAAAACCTGTTCCCGCGGGGAGAAATTTTCTCTCCCGGCC 1740
Db 1681 TGGGGGAAAAACAGGGGTAAAAACCTGTTCCCGCGGGGAGAAATTTTCTCTCCCGGCC 1740

Qy 1741 TCAATATTTCCCGCACCAACCAATCATGAGCGGGCGCCCGGAAAGCATATAAGTGT 1800
Db 1741 TCAATATTTCCCGCACCAACCAATCATGAGCGGGCGCCCGGAAAGCATATAAGTGT 1800

(META-) METAGEN GES GENOMFORSCHUNG MBH.
Rosenthal A, Hinzmann B, Schaefer R, Zuber J, Tchernitsa O;
Grips M, Hellriegel M, Schmitz A, Sers C;
WPI; 2001-483415/52.
Nucleic acids differentially expressed between tumor and normal cells,
useful for diagnosis or therapy of tumors and for screening active
agents
Disclosure; Page 362; 579pp; German.
This invention describes a nucleic acid (I) with differential expression
between tumour and normal cells and which has cytostatic activity. (I)
work as modulators of Ras activity by inducing expression of tumour
suppressor genes. (I), and polypeptides encoded by them, are useful as
targets for diagnosis or therapy and in screening to determine the
effects of an active compound (potential pharmaceutical) on a cell line,
particularly for diagnosis and treatment of tumors, especially by
modulating expression of (I) (by gene therapy, antisense RNA or ribozyme
methods) or by modulating the amount and/or location of (I)-encoded
polypeptides (by administration of the polypeptide or its activator,
antibody (optionally as a conjugate) or inhibitor). The method allows
identification of many Class II tumour suppressor genes (i.e. genes that
are not primary targets for tumour-initiating mutations).
AAH81492-AAH82376 represent the human and rat derived nucleic acid
fragments described in the method of the invention.
Sequence 318 BP; 53 A; 97 C; 89 G; 69 T; 0 other;
Query Match 15.0%; Score 296.4; DB 22; Length 318;
Best Local Similarity 98.0%; Pred. No. 7.3e-76;
Matches 300; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 130 GGACGCAAGTACATCCAGACAGACAGCGGCCCTACTGTGCTGCTATGACAAATACC 189
DB 2 GGACGCAAGTACATCCAGACAGACAGCGGCCCTACTGTGCTGCTATGACAAATACC 61
QY 190 TTGTCACACACCTGTGCTGAGTGCCAGCAGCTTATCGGCGCAGTCTGAGGAGCTGTC 249
DB 62 TTGTCACACACCTGTGCTGAGTGCCAGCAGCTTATCGGCGCAGTCTGAGGAGCTGTC 121
QY 250 TATGAAGACCGCCATTTCCAGAGGGCTGCTCCGCTGCTGCGCTGCGAGCGCTCACTA 309
DB 122 TATGAAGACCGCCATTTCCAGAGGGCTGCTCCGCTGCTGCGCTGCGAGCGCTCACTA 181
QY 310 GCCGATGACCTTACCTGCCAGCAGTACGAGTCTGCTGCAATGACTGCTACTGCACT 369
DB 182 GCCGATGACCTTACCTGCCAGCAGTACGAGTCTGCTGCAATGACTGCTACTGCACT 241
QY 370 CGGTTTCTCCTGCAGTCTCGCTTGTGGGAGAGCTGTCATGCTGCTGCGGAGCTG 429
DB 242 CGGTTTCTCCTGCAGTCTCGCTTGTGGGAGAGCTGTCATGCTGCTGCGGAGCTG 301
QY 430 GAATAT 435
DB 302 GGAAT 307
RESULT 4
ID ABQ99467
XX ABQ99467 standard; cDNA; 1158 BP.
AC ABQ99467;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human coding sequence SEQ ID 200.
XX
KW Human; expressed sequence tag; EST; chromosome 6q16.1-q16.3;
KW haematopoietic disorder; central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;

immune deficiency; immune disorder; bacterial infection; allergy; cancer;
fungal infection; autoimmune disorder; coagulation disorder; neotropic;
antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
cytostatic; haemostatic; virucide; antibacterial; fungicide;
immunostimulant; cerebroprotective; gene therapy; gene; ss.
Homo sapiens.
WO200259260-A2.
01-AUG-2002.
16-NOV-2001; 2001WO-USA2950.
17-NOV-2000; 2000US-0714936.
(HYSE-) HYSEQ INC.
Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
WPI; 2002-590824/63.
N-PSDB; ABP64881.
New isolated polynucleotide, useful in research, diagnostic or
therapeutic methods, e.g. preventing or treating disorders involving
aberrant protein expression or biological activity -
Claim 1; SEQ ID 200; 394pp; English.
The present invention relates to novel human coding sequences
(ABQ99268-ABQ99608) and proteins (ABP64682-ABP65022). The sequences are
useful in therapeutic, diagnostic and research methods. The
polynucleotides may be used in the field of molecular biology as
hybridisation probes, primers for PCR, for chromosome and gene mapping,
for the recombinant production of protein, or in generation of anti-sense
DNA or RNA. The polynucleotides are useful in diagnostics as expressed
sequence tags (ESTs) for identifying expressed genes or for physical
mapping of the human genome. The proteins may be used as molecular weight
markers, or as nutritional sources or supplements. The proteins may be
used to maintain and expand cell population in a totipotent or
pluripotent state useful for re-engineering damaged or diseased
tissues, transplantation, manufacture of bio-pharmaceuticals or the
development of bio-sensors. The polynucleotides and proteins are useful
for preventing, treating or ameliorating disorders involving aberrant
protein expression or biological activity, e.g. haematopoietic disorders,
central/peripheral nervous system diseases, mechanical and traumatic
disorders, non-healing wounds, immune deficiencies and disorders,
infectious diseases caused by viral, bacterial or fungal infection,
autoimmune disorders, allergic reactions and conditions, coagulation
disorders, or cancer. The polynucleotide sequences of the invention were
assembled from ESTs isolated mainly by sequencing by hybridisation, and
in some cases, sequences obtained from one or more public databases.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
Sequence 1158 BP; 337 A; 238 C; 263 G; 320 T; 0 other;
Query Match 13.3%; Score 262.4; DB 24; Length 1158;
Best Local Similarity 57.6%; Pred. No. 1.1e-65;
Matches 470; Conservative 0; Mismatches 346; Indels 0; Gaps 0;
QY 97 TTGACTGTGCAAAATGCAACGAGTCCTGTATGACGCAAGTACATCCAGACACAGC 156
DB 247 TTTTACTGTCAATCTGCACAGCATCTCTTGGGAGAGAATATGTACTAAGGATGAC 306
QY 157 GCGCCCTACTGTGTGCCCTGCTATGACAAATACCTTTGCCAACACCTGTGCTGAGTCCAG 216
DB 307 AGTCCATCTGTGTACATGTTATGATCGTGTATTTTCTTAACATTTGCGAGGAATGCAA 366
QY 217 CAGCTTATCGGGCATGACTCGAGGAGCTGTTCATGAGACCCGCTTCATCCAGGCGC 276
XX
XX

DB 906 AGCGTTTACCGCTCTGGAGGACCAAGTATTACTGGCTGGATTGCTACAGAACTTTGGG 965
QY 737 CACCTAAGTGCAGCAGCTGCAAGCGCCCATCTGCTAGGACTCGGTGGAGS---CAAGTATG 793
DB 966 CCAAGAAGTGTCTGGATGCAAGAACCACCATCACTGGGTTGGTAAAGGCTCCAGTGGG 1025
QY 794 TGTCTTTTGAAGACCGACACTGGCACCACAACTGCTTCTGCGCCGCTGCTCTAAGCT 853
DB 1026 TGGCCTATGAAGGACAATCTTGGCAGCACTACTGCTTCCACTGCAAAAATGCTCCGTA 1085
QY 854 CCCTGGTGGGAGGGCTTGTACCGGATGAGACCAAGTCTGTGCGAGGCTGTAGCC 913
DB 1086 ATCTGGCCCAACAAGCGCTTTGTTTCCACGAGGAGAGTGTATTGTCCCGACTGTGCA 1145
QY 914 A 914
DB 1146 A 1146

RESULT 6

ABK34596/c

ID ABK34596 standard; cDNA; 2340 BP.

AC ABK34596;

DT 08-MAY-2002 (first entry)

DE Human cDNA for novel secreted protein, SEQ ID 365.

XX Human; ss; gene; secreted protein; immune deficiency; viral infection;
KW bacterial infection; fungal infection; autoimmune disorder; burn;
KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
KW coagulation disorder; haemophilia; inflammatory disorder; ulcer;
KW tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
KW lymphoid cell deficiency.

XX Homo sapiens.

OS WO200177290-A2.

PN 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US10295.

PF 06-APR-2000; 2000US-194941P.

PR (GENY) GENETICS INST INC.

XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;

PI Gulukota K, Graham JR;

XX WPI; 2002-179323/23.

PT Six hundred and twenty five polynucleotides derived from a variety of
PT human tissue sources which encode secreted proteins, useful for
PT treating immune deficiencies and disorders such as autoimmune disorders

XX Claim 1; Page 183-184; 339pp; English.

XX The invention relates to 625 polynucleotides which have been derived from
CC a variety of human tissue sources and which encode novel secreted
CC proteins, their complements and sequences that hybridise to them.
CC Also included are a vector comprising the polynucleotide, a host cell
CC transformed with the vector, the proteins encoded by the
CC polynucleotides, antibodies that bind to the proteins and identification
CC of modulators of the proteins or the expression of the polynucleotide.
CC The polynucleotides can be used as probes for the identification
CC and isolation of full length cDNA and genomic DNA. The polynucleotides
CC and proteins can also be used as nutritional supplements. The protein

CC is useful in the treatment of various immune deficiencies and disorders
CC such as viral infections, bacterial infections, fungal infections,
CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,
CC autoimmune thyroiditis and diabetes) and allergic reactions and
CC conditions (e.g. asthma). They are also useful for treating
CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
CC useful for tissue regeneration, for wound healing and in the treatment
CC of burns, incisions and ulcers. The proteins are also useful for
CC regulating haematopoiesis, for treating myeloid or lymphoid cell
CC deficiencies. The present sequence is one of the 625 cDNA sequences
XX encoding a secreted protein.

QY Sequence 2340 BP; 621 A; 515 C; 554 G; 650 T; 0 other;

Query Match 13.0%; Score 255.4; DB 24; Length 2340;
Best Local Similarity 57.3%; Pred. No. 1.7e-63;
Matches 482; Conservative 0; Mismatches 356; Indels 3; Gaps 1;

QY 77 TTGCCACCATGAGCGAGTCATTGTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGA 136
DB 2210 TGGGCACCATGGCGGAGAGTTTGTACTGCTACTGAGGATCCCTTGCAGGGAAGA 2151
QY 137 AGTACATCCAGACAGACAGCGCCCTACTGTGTGCTGCTATGACAAATACCTTTGCCA 196
DB 2150 AGTATGTCAAAAGGATGCCACCACTGCTGCTGAAATGCTTTGACAAGTTCTGTGCCA 2091
QY 197 ACACCTGTGTGAGTGCCAGCAGCTATCGGCATGCTCGAGGAGCTGTTCTATGAAG 256
DB 2090 ACACCTGTGTGAAATGCCGCAAGCCATCGGTGCGGACTCCCAAGGAGGTGCACTATAAGA 2031
QY 257 ACCGCGATTTCACGAGGCTGCTTCGCTGCTGCTGCCAGCGCTCCTACCTAGCCGATG 316
DB 2030 ACCGCTTCTGGCATGACACCTGCTTCGCTGCTGCAAGTGCCTTCCACCTTTGGCCAAATG 1971
QY 317 AACCTTCACCTGCCAGGACAGTGTGCTGCTGCAATGCTGTCTACTGTGCTGCTGCTTTT 376
DB 1970 AGACCTTTGTGCCAAGGACAAAGATCCTGTGCAACAAGTGCACTCTGGGAGGACT 1911
QY 377 CCTGCGAGTCTCCCTTTGTGGGAGACTGTCTATCCCTGGGTCCCGGAGCTGGGAATATG 436
DB 1910 CCCCCAAGTGAAGGGGTGCTTCAAGGCCATTTGTGCGAGGAGATCAAAACGTTGGATACA 1851
QY 437 GAGGCCAGACATGGCATGAGCACTCTTCTCTGTCAGTGGCTGTCAACAGCCACTGGGCT 496
DB 1850 AGGGACCGCTTGGCAAAAGACTGCTTCACTGTAGTAAGTCAAGCAAGTCACTCGGA 1791
QY 497 CCGCTTCTTTTGTGCCGACAAAGGTGCTCACTACTGCTGCTGCTGCTGCTATGAGAAAGT 556
DB 1790 CTGGAAGCTTCTTCCCTAAAGGGGAGGACTTCTACTGCTGCTGCTGCTGCTGCTGCTGCT 1731
QY 557 TTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
DB 1730 TTGCCAAGCATTGGCTGAAAGTGCACAAAGGCGCATCACATCTGGAGGAATCACTTACCAGG 1671
QY 617 ATGAGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
DB 1670 ATGAGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1611
QY 677 AGCAGTTTCCACCTCCCGGATGAAGATCCCTACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
DB 1610 AGCGTTTCCACCTGCTGGAGGACCAAGTATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1551
QY 737 CACCTAAGTGCAGCAGCTGCAAGCGCCCATCTGAGGACTCGGTGGAGS---CAAGTATG 793
DB 1550 CCAAGAAGTGTGCTGGATGCAAGAACCCCATCACTGCGTTTGGTAAAGGCTCCAGTGTGG 1491
QY 794 TGTCTTTTGAAGACCGACACTGGCACCACCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853
DB 1490 TGGCCTATGAGGACAACTCTTGGCAGCACTACTGCTTCCACTGCAAAAATGCTCCGTA 1431
QY 854 CCCTGGTGGCGCAGGGCTTCTGATACGGATGGAGACCAAGTGTCTGTGCCAGGCTGTAGCC 913

Db 1430 ATCTGCCCAACAGCGCTTGTGTTTCCACCAGGAGCAAGTGATTGTCGCCGACTGGGCCA 1371
QY 914 A 914
Db 1370 A 1370

RESULT 7
ABL62451
ID ABL62451 standard; DNA: 2254 BP.

AC ABL62451;
XX
DT 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:788.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX

XX 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-231133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 28-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237117P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.

XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 788; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 2254 BP; 611 A; 537 C; 498 G; 608 T; 0 other;

Query Match 12.8%; Score 252.2; DB 24; Length 2254;
Best Local Similarity 57.1%; Pred. No. 1.4e-62;
Matches 480; Conservative 0; Mismatches 358; Indels 3; Gaps 1;
QY 77 TTGCCACCATGAGCGAGTCAATTTGACTGTGCAAAATGCAACGAGTCCCTGTATGGAGCA 136
Db 76 TGGGCACCATGGCGGAGAGATTTCAGTCCACTACTGCGAGGATCCCTTCAGGGGAGAGA 135
QY 137 AGTACATCCAGACAGACAGCGCCCTCTACTGTGTGCTGCTCTATGACAAATACCTTTGCCA 196
Db 136 AGTATGTGCAAAAGGATGGCCACCACTGCTGCTGCTGAAATGCTTTGACAAGTTCTGTGCCA 195
QY 197 ACACCTGTGCTGAGTGGCCAGCAGCTTATCGGGGATGACTCGAGGGAGCTGTTATGTAAG 256
Db 196 ACACCTGTGTTGGAATCCGCAAGCCCATCGGTGCGGACTCCAAGGAGGTGCATATAAGA 255
QY 257 ACCGCCATTTCCACAGAGGCTGCTCCGCTGTGCTGCGCTGCCAGCGCTACTAGCCGATG 316
Db 256 ACCGCTTCTGGCATGACACCTGCTTCCGCTGTGCTGCTGCTGCTTCAACCCCTTGGCCATG- 315
QY 317 AACCTTCACTGCCAGGAGTGAAGTGTCTGTGCAATGACTGCTACTGCAAGTGCCTTTT 376
Db 316 AGACCTTTGTGGCCAAAGGACAAAGATCCTGTGCAACAGTGCACCACTCGGAGGACT 375
QY 377 CCTCGCAGTGTCCGCTTGTGGGGAGACGTGCTATGCTGGTCCCGGAAGCTGGAATATG 436
Db 376 TCCCAAGTGCAAAGGGGTGCTTCAAGGCCATTTGTCAGGAGATCAAAAGCTGAGTACA 435
QY 437 GAGGCCAGACATGGCATGAGCACTGCTTCTGTGTCAGTGTGCTGTGTCAGCAGCTGGCT 496
Db 436 AGGGGACCGTCTTGGCACAAAGACTGCTTACCTGTGTAGTACTCAAGCAAGTCAATCGGGA 495
QY 497 CCGCTTCTTTTGTGCCCGACAAAGGTGCTCACTACTGCTGCTGCTGCTATGAGAACAAAGT 556
Db 496 CTGGAAGCTTCTTCCCTAAGGGGAGGAGCTTCTACTGCTGCTGCTGCTGCTGCTGCTGCT 555
QY 557 TTGCTCCTCGTGCGCCCGCTGCAACGAAGCGTGCACAGGCTGAGAGGTGAGTACATACCGTG 616
Db 556 TGGCCCAAGCATTCGCTGGAAGTGCAACAAGGCCCATCATCTCTGGAGGAATCACTTACCAGG 615

DT 21-NOV-2001 (first entry)
XX Human reproductive system related antigen cDNA SEQ ID NO: 1354.
DE Human; reproductive system related antigen; reproductive system disorder;
XX cancer; gene therapy; ss.
KW Homo sapiens.
KW WO200155320-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01339.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231422.
PR 08-SEP-2000; 2000US-0231423.
PR 08-SEP-2000; 2000US-0231424.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-02331968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.

```
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-465570/50.
XX P-PSDB; AAM95383.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen
XX is used in preventing, treating or ameliorating a medical condition -
XX
XX Claim 1; SEQ ID NO 1354; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a coding sequence of the
XX invention.
XX
XX Sequence 830 BP; 234 A; 181 C; 197 G; 212 T; 6 other;
XX
XX Query Match 11.8%; Score 232.6; DB 22; Length 830;
XX Best Local Similarity 57.5%; Pred. No. 4.4e-57;
XX Matches 409; Conservative 4; Mismatches 298; Indels 0; Gaps 0;
XX
QY 202 TGTGCTGAGTGCACAGCTTATTCGGGATGACTCGAGGGAGCTGTTCTATGAAGACCGC 261
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4 TCGAGGAATGCAAAAACCAATTAATCTGATTAAGGATCTTTGTACAAAGACGG 63
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 262 CATTTCCAGGGCTGCTCGCTGCTGCCCTGCCAGCGCTCACTAGCCGATGAACC 321
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 64 CACTGGCATGAAGGATGCTTAAGTGACCAATGCAATCACTCTTTGGTGGAAAGCCT 123
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 322 TTCACCTGCCAGCAGTGAGTCTGCTGCAATGACTGCTACTGCAGTCCGTTTTCCTCG 381
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 124 TTTGCTGCCAAGGATGAGCGCTGCTGTGCAGGAGTCTATCTAATGAGTCTCTCC 183
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 382 CAGTGCCTCGTGTGGGAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 441
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 184 AAGTGTCTTCCACTGAAGGAGGACCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 442 CAGACATGSCATGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 244 AACTACTGCGATGAWACCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 303
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 502 TCTTTTGTGCCGCAAGGCTGCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 304 CTTTGTGATCTCAAGAGAGTGGCAATATTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 363
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 562 COTCCTGCTGCGCGCTGCGAGCAAGAGCTGTACAGAGGTGGAGTGACATACCGTGATCAG 621
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 364 CACTACTGCACTTTTGTGAAGAAGGTGATACTTCAAGTGGGATAAATTTTGTGACCAAG 423
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 622 CCGTGGCATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 681
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 424 CTATGGCATAAAGATGTTTCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 483
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 682 TTCACCTCCGGGATGAAGATCCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 484 TTCATGTCAGAGAGACTATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 742 AAGTGCAGCAGTGCAGGCGCCCATCTAGTACTCGGTGGAGGCAAGTATGTGCTCTTT 801
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 544 AAGTGTGATGCTGTTTCAACCCCATTAAGTGTCTCAGAGTGGCCAACTTATCTGCTTT 603
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 802 GAAGACGACACTGCGACACCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 861
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 604 CAAGACAGCCAGTGGCATAGCGAATGCTTTAACTGCGGGAAATGCTCTGCTCTCTTGGTG 663
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 862 GCCCAGGGCTGTCTACCGGATGGAGACCAAGTGTCTTCCAGGGCTGTAGC 912
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 664 GGTAAAGGCTTCTGACCCAGAACAGGAATCTTCTGCCAAATGTGGC 714
XX
XX RESULT 12
XX AAS33146
XX ID AAS33146 standard; CDNA; 830 BP.
XX
XX AC AAS33146;
XX
XX DT 04-DEC-2001 (first entry)
XX
XX DE DNA encoding human secreted protein, Seq ID No 105.
XX
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
XX rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;
XX cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
XX cytostatic; Alzheimer's disease; Parkinson's disease; human cancer;
XX multiple sclerosis; cancer; hyperproliferative disorder; infection;
XX Gaucher's disease; neurological disease; cerebrovascular disorder;
XX thrombosis; wound healing; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200155326-A2.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US01347.
XX
XX PR 31-JAN-2000; 2000US-0179065.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX
XX DR WPI: 2001-451931/48.
XX
XX DR P-PSDB; AAU20437.
XX
XX PT New nucleic acids and polypeptides, useful for diagnosing, preventing
XX or treating medical conditions -
XX
XX PS Claim 1; SEQ ID No 105; 753pp; English.
XX
XX CC The invention relates to novel isolated nucleic acid molecules (I)
XX encoding human secreted proteins (II). (I) and (II) are used to prevent,
XX treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
XX goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
XX the prevention, treatment and diagnosis of diseases associated with
XX inappropriate expression of secreted proteins. (I) and complementary
XX sequences may also be used as DNA probes in diagnostic assays (e.g.
XX polymerase chain reactions (PCR)) to detect and quantitate the presence
XX of similar nucleic acid sequences in samples, and so which patients may
XX be in need of restorative therapy. (II) may also be used as antigens in
XX the production of antibodies and in assays to identify modulators
XX (agonists and antagonists) of the expression and activity of the secreted
XX proteins. The anti-(II) antibodies and antagonists may also be used to
XX down regulate expression and activity of (II). The anti-(II) antibodies
XX may also be used as diagnostic agents for detecting the presence of (II)
XX in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
XX disorders include for example: immune/autoimmune diseases (e.g. HIV
XX (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
XX and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
XX melanomas, neoplasms of the breast or liver, Sezary syndrome and
XX Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
XX Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
XX cerebrovascular disorders (e.g. cardiac arrest, tachycardia,
XX angina and thrombosis), infections caused by bacteria, viruses and
XX fungi and ocular disorders (e.g. corneal infections). (I) and (II),
XX agonists, antagonists and antibodies can also be used to promote wound
XX healing, maintain organs before transplantation, and support cell culture
```


22
127 ncccccttctgttttcaccccccattccttgcctgctccatcatatgaaga 250

Matches	472;	Conservative	0;	Mismatches	366;	Indels	3;	Gaps	1;
Qy	77	TTGCCACCATGAGCGAGTCATTTTACACTGTGTGCAAAATGCAACGAGTCCCTGTATGGACGCA	136						
Db	76	TGGCACCATTGGCGGAGAAGTTTGACTGCCACTACTGCAGGGATCCCTTCGACGGGGAAGA	135						
Qy	137	AGTACATCCAGACAGACAGCGGGCCCCPACTGTGTGCGCCCTGCTATGACAATACCTTTTGCCA	196						
Db	136	AGTATGTGCAAAAGGATGGCAGCACACTGCTGCCTGAAATGCTTTGACAAGTCTGTGTGCCA	195						
Qy	197	ACACCTGTGCTGAGTGCACGACGCTTATCGGGCATGACTCGAGGGAGCTGTTCTATGAAG	256						
Db	196	ACACCTGTGTGAATGGCGAAGCCCATCGGTGGGACTCCAAGAGGTGCATATATAAGA	255						
Qy	257	ACCGCCATTTCACAGAGGGCTGCTTCGCGTGTGCCGTGCCAGCGCTCACTAGCCGATG	316						
Db	256	ACCGCTTCTGGCATGACACCTGCTTCGCGTGTGCCAAGTGCCCTTCACCCCTTGCCCAATG	315						
Qy	317	AACCTTTCACCTGCCAGGACAGTGAAGTGTCTGCAATGACTGCTACTGCAGTGCCTTTT	376						
Db	316	AGACCTTTTGTGGCCAAAGGACAACAAGATCCCTGTGCACAGTGCACCATNGGAGGACT	375						
Qy	377	CCTCGCAGTCTCGCTTGTGGGAGACTCTCATGCTGGTGGGCCGGAAGCTGGAATATG	436						
Db	376	TCGCCAAGTGCAAGGGGTGCTTCAAGGCCATTGTGGCAGGAGATCAAAAGCTGGAGTACA	435						
Qy	437	GAGGCCAGACATGGCATGACACTGCTTCCTGTGTCAGTGGCTGTGAACGCCACTGGGCT	496						
Db	436	AGGGACCGTCTGGCACAAGACTGCTTCACTGTAGTAGTACTGCACGCAAGTATCGGGA	495						
Qy	497	CCCGTTCTTTTGTGCCGACAAGGGTGCTCACTACTGCGTGCCTGCTGCTATGAGAACAAAT	556						
Db	496	CTGGAAGCTTCTTCCCTAAGGGGAGGACTTCTACTGCGTGACTTGCGCATGAGACCAAGT	555						
Qy	557	TTGCTCTCTCGTGGCCCGCTGCACGAAGACGCTGCACAGGGTGGAGTGACATACCGTG	616						
Db	556	TGGCCAAAGCATTGGCTGAAGTGAACAAGGCCATCATCTGAGGGAATCACATTACCAAG	615						
Qy	617	ATACGCCGTGGCATCGAAGATGTCTGTGCTGTACCGGATCCAGACGCCCTGGCAGGGC	676						
Db	616	ATCAGCCCTGGCATGCCGATTTGCTTTGTGTGTGTACCTGCTCTAAGAACTGGCTGGGC	675						
Qy	677	AGCAGTTTCACTCCCGGGATGAAGATCCCTACTGTGTGGCCTGTTTGGAGAACTCTTTG	736						
Db	676	AGCGTTTTCACCGCTGTGGAGGACAGTATTACTGCGTGGATTGCTCAAGAACTCTGTGG	735						
Qy	737	CACCTAAGTGCACGAGCTGCAAGGCCCCCATCGTAGGACTCGGTGGAGG---CAAGTATG	793						
Db	736	CCAAGAAGTGTGTGAATGCAAGAACCCCCATCACTGGGTTTGGTAAAGGCTCCAGTGTGG	795						
Qy	794	TGTCCTTTGAAGACCGACACTGGCACCACAACTGCTTCTCTCGGCCCGCTGCTACCT	853						
Db	796	TGGCCTATGAAGNCAATCTCTGGCAGCACTACTGCTTCCACTGCAGAAAATGCTTCGTGA	855						
Qy	854	CCCTGGTGGCCAGGGCTTCGTATCCGGATGGAGCAACCAAGTGCCTCGCAGGGCTGTAGCC	913						
Db	856	ATCTGGCCCAACAAGCGCTTGTGTTTCCACACGAGCAAGTGTATGTTCOCGACTGTGCCA	915						
Qy	914	A	914						
Db	916	A	916						

? TITLE OF INVENTION: No. 6444803el Bone Mineralization Proteins, DNA, Vectors
 ?
 ? FILE REFERENCE: 06148.0115
 ? CURRENT APPLICATION NUMBER: US/09/721,975
 ? CURRENT FILING DATE: 2000-11-27
 ? PRIOR APPLICATION NUMBER: US 05/124,238
 ? PRIOR FILING DATE: 1998-07-29
 ? PRIOR APPLICATION NUMBER: 60/054,219
 ? PRIOR FILING DATE: 1997-07-30
 ? PRIOR APPLICATION NUMBER: 60/080,407
 ? PRIOR FILING DATE: 1998-04-02

RESULT 2
US-09-124-238A-8
; Sequence 8, Application US/09124238A
; Patent No. 6300127
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6300127el Bone Mineralization Proteins, DNA, Vectors
; TITLE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1597 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1362
US-08-166-316-1

Query Match 3.8%; Score 75.4; DB 1; Length 1597;
Best Local Similarity 48.7%; Pred. No. 9.6e-12;
Matches 239; Conservative 0; Mismatches 246; Indels 6; Gaps 1;
QY 416 GGTCCCGGAAGCTGGAATATGGAGCCGACATGGAGCCGACATGCATGCTTCCGTGCAGTG 475
DB 860 GGGGCGGTACCTGGTGGCGTGGCGCCAGCGTACCCCGGAGGAGTTGTGTAGGCC 919
QY 476 GCTGTGAACAGCACTGGGCTCCCTTCTTTGTGCCGACAAAGGGTGTCTACTACTGCG 535
DB 920 AGTGTGGGAAGTCTCTGGAAGAGGTGGCTTCTTTGAGGAGAAGGGCGCCATCTTCTGCC 979
QY 536 TCCCTGTCTATGAGAACAAAGTTTGTCTCTCGCTGGCGCGCTGACAGACAGCGCTGACAC 595
DB 980 CACCATGTATGACGTGGCTATGACCCAGCTGTGCCAAGTGTGCAAGAAGAAGATTACAG 1039
QY 596 AGGGTGGAGTGCATACCGTGTATCAGCCGTGGCATCGAGAAATGCTGTGCTACCGGAT 655
DB 1040 GCGAGATCATGACCCCTGAAGATGACCTGGCAGTGCACCTGCTTACCTGTGCTGCT 1099
QY 656 GCCAGAGCCCTTGGCAGGCGACAGTTCACCTCCCGGATGAAGATCCCTACTGTGTGG 715
DB 1100 GCAAGACGCCATCCGGAAGAGGGCTTCTACATGAGGAGGGCGTGGCCCTATTGGAGC 1159
QY 716 CTTGTTTGGAGAACTTTTGCACCTAAGTGCAGAGCTGCAAGCGCCCGCATCGTAGGAC 775
DB 1160 GAGACTATGAGAAGATGTTTGGCAGAAATGCCATGGTGTGACTTCAAGATCG----- 1213
QY 776 TCGGTGGAGGCAAGTATGTCTTTTGAAGACCGACATGGCACCACCACTGCTTCTCT 835
DB 1214 AGCTGGGAGCGCTTCTTGGAGGCCCTTGGCTTTCAGCTGGCATGACACCTGCTTCTCT 1273
QY 836 GCGCCCGCTGTCTTACCTCCCTGGTGGCCAGGGCTTCTACCCGGATGAGACCAAGTGC 895
DB 1274 GTGCGATATGTCAGATCAACCTGGAAGGAAGAACCTTCTACTCCAAGAGGACAGGCCCTC 1333
QY 896 TCTGCCAGGCG 906
DB 1334 TCTGCAAGAGC 1344

RESULT 6
US-09-124-238A-32
; Sequence 32, Application US/09124238A
; Patent No. 6300127
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6300127el Bone Mineralization Proteins, DNA, Vectors,
; FILE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/124,238A
; CURRENT FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 32
; LENGTH: 1620

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-124-238A-32

Query Match 3.8%; Score 75.4; DB 4; Length 1620;
Best Local Similarity 48.7%; Pred. No. 9.7e-12;
Matches 239; Conservative 0; Mismatches 246; Indels 6; Gaps 1;
QY 416 GGTCCCGGAAGCTGGAATATGGAGCCGACATGGAGCCGACATGCATGCTTCCGTGCAGTG 475
DB 869 GGGGCGGTACCTGGTGGCGTGGCGCCAGCGTACCCCGGAGGAGTTGTGTAGGCC 928
QY 476 GCTGTGAACAGCACTGGGCTCCCTTCTTTGTGCCGACAAAGGGTGTCTACTACTGCG 535
DB 929 AGTGTGGGAAGTCTCTGGAAGAGGTGGCTTCTTTGAGGAGAAGGGCGCCATCTTCTGCC 988
QY 536 TCCCTGTCTATGAGAACAAAGTTTGTCTCTCGCTGGCGCGCTGACAGACAGCGCTGACAC 595
DB 989 CACCATGTATGACGTGGCTATGACCCAGCTGTGCCAAGTGTGCAAGAAGAAGATTACAG 1048
QY 596 AGGGTGGAGTGCATACCGTGTATCAGCCGTGGCATCGAGAAATGCTGTGCTTACCGGAT 655
DB 1049 GCGAGATCATGACCCCTGAAGATGACCTGGCAGTGCACCTGCTTACCTGTGCTGCT 1108
QY 656 GCCAGAGCCCTTGGCAGGCGACAGTTCACCTCCCGGATGAAGATCCCTACTGTGTGG 715
DB 1109 GCAAGACGCCATCCGGAAGAGGGCTTCTACATGAGGAGGGCGCTGCGCTATTGGAGC 1168
QY 716 CTTGTTTGGAGAACTTTTGCACCTAAGTGCAGAGCTGCAAGCGCCCGCATCGTAGGAC 775
DB 1169 GAGACTATGAGAAGATGTTTGGCAGAAATGCCATGGCTGTGACTTCAAGATCG----- 1222
QY 776 TCGGTGGAGGCAAGTATGTCTTTTGAAGACCGACATGGCACCACCACTGCTTCTCTCT 835
DB 1223 AGCTGGGAGCGCTTCTTGGAGGCCCTTGGCTTTCAGCTGGCATGACACCTGCTTCTCT 1282
QY 836 GCGCCCGCTGTCTTACCTCCCTGGTGGCCAGGGCTTCTACCCGGATGAGACCAAGTGC 895
DB 1283 GTGCGATATGTCAGATCAACCTGGAAGGAAGAACCTTCTACTCCAAGAGGACAGGCCCTC 1342
QY 896 TCTGCCAGGCG 906
DB 1343 TCTGCAAGAGC 1353

RESULT 7
US-09-721-975-32
; Sequence 32, Application US/09721975
; Patent No. 644803
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 644803el Bone Mineralization Proteins, DNA, Vectors,
; FILE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/721,975
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/124,238
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 32
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-721-975-32

Query Match 3.8%; Score 75.4; DB 4; Length 1620;
Best Local Similarity 48.7%; Pred. No. 9.7e-12;

Db 1049 GCGAGATCATGACGGCCCTGAAGATGACCTGGCAGCTGCACCTTTACCTGTGCTGCT 1108
QY 656 GCCAGACGCCCTGGCAGGCGACAGTTTCACCTCCCGGATGAAGATCCCTACTGTGTGG 715
Db 1109 GCAAGAGCCCATCCGGAACAGGCTTCTACATGGAGGAGGCGCTGCTTATTCGGAGC 1168
QY 716 CCGTGTGGAGAACTCTTTGGACCTAAGTGCAGCAGCTGCAAGCGCCCATCGTAGGAC 775
Db 1169 GAGACTATGAGAAGATGTTGGCAGCAATGCTGCTGCTGCTTCAAGATCG----- 1222
QY 776 TCGGTGGAGCAAGTATGTGCTTCTTGAAGCCGACACTGGCACCACCAACTCTCTCT 835
Db 1223 ACGTGGGAGCGCTTCCTGGAGGCCCTGGGCTTCAGCTGGCATGACACCTGCTTCGTCT 1282
QY 836 GGGCCCGCTGCTTACCTCCCTGGTGGCCAGGGCTTCGTACCGGATGGAGACCAAGTGC 895
Db 1283 GTGCGATATGTCAGATCAACCTTGAAGGAAAGACCTTCTACTCCAAAGAGGACGGCCTC 1342
QY 896 TCTGCCAGGC 906
Db 1343 TCTGCAAGAGC 1353

RESULT 10

US-09-721-975-9
; Sequence 9, Application US/09721975
; Patent No. 6444803
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6444803el Bone Mineralization Proteins, DNA, Vectors,
; FILE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/721,975
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/124,238
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 9
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-721-975-9

Query Match 3.8%; Score 75.4; DB 4; Length 1644;
Best Local Similarity 48.7%; Pred. No. 9.8e-12;
Matches 239; Conservative 0; Mismatches 246; Indels 6; Gaps 1;
QY 416 GGTCCCGAAGCTGGAATATGGAGCCAGACATGCGATGACGACTGCTTCTGTGCACTG 475
Db 869 GGGCCCGTACTGTGTGGCTTGGGCCACGCGTACACCCCGAGGAGTTGTGTGTAGCC 928
QY 476 GCTGTGAACAGCACTGGCTCCCGTCTTTTGTGCCGACAAAGGCTGCTACTACTGCG 535
Db 929 AGTGTGGAGGTCCTGGAAGGGTGGCTTCTTTGAGGAGAAGGCGCCATCTTCTGCC 988
QY 536 TGCCCTGTATGAGAACAAAGTTTGTCTCTCGCTGGCGCGCTGCGAGCAAGCGCTGACAC 595
Db 989 CACCATGTATGACGTGGCTATGCACCCAGCTGTGCAAGTGCACCAAGAAGATTACAG 1048
QY 596 AGGTGGAGTGCATACCTGATCAGCGTGCATCGAGATGCTGTCTCTACCGGAT 655
Db 1049 -GCGAGATCATGACGCCCTGAAGATGACCTGGCAGCTGCACCTTTACCTGTGCTCT 1108
QY 656 GCCAGACGCCCTGGCAGGCGACAGTTTCACCTCCCGGATGAAGATCCCTACTGTGTGG 715
Db 1109 GCAAGACCCCATCCGAACAGGCGCTTCTACATGGAGGAGGCGCTGCTTATTCGGAGC 1168
QY 596 AGGTGGAGTGCATACCTGATCAGCGTGGCATCGAGATGCTGTCTGTACCGGAT 655
Db 1049 GCGAGATCATGACGCCCTGAAGATGACCTGGCAGCTGCACCTTTACCTGTGCTCT 1108
QY 656 GCCAGACGCCCTGGCAGGCGACAGTTTCACCTCCCGGATGAAGATCCCTACTGTGTGG 715
Db 1109 GCAAGACCCCATCCGGAACAGGCGCTTCTACATGGAGGAGGCGCTGCTTATTCGGAGC 1168

QY 716 CCGTGTGGAGAACTCTTTGGACCTAAGTGCAGCAGCTGCAAGCGCCCATCGTAGGAC 775
Db 1169 GAGACTATGAGAAGATGTTGGCAGCAATGCTGCTGCTTCAAGATCG----- 1222
QY 776 TCGGTGGAGCAAGTATGTGCTTCTTGAAGCCGACACTGGCACCACCAACTCTCTCTCT 835
Db 1223 ACGTGGGAGCGCTTCCTGGAGGCCCTGGGCTTCAGCTGGCATGACACCTGCTTCGTCT 1282
QY 836 GGGCCCGCTGCTTACCTCCCTGGTGGCCAGGGCTTCGTACCGGATGGAGACCAAGTGC 895
Db 1283 GTGCGATATGTCAGATCAACCTTGAAGGAAAGACCTTCTACTCCAAAGAGGACGGCCTC 1342
QY 896 TCTGCCAGGC 906
Db 1343 TCTGCAAGAGC 1353

RESULT 11

US-09-986-621-9
; Sequence 9, Application US/09986621
; Patent No. 6521750
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6521750el Bone Mineralization Proteins, DNA, Vectors,
; FILE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/986,621
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 09/124,238
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 9
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-986-621-9

Query Match 3.8%; Score 75.4; DB 4; Length 1644;
Best Local Similarity 48.7%; Pred. No. 9.8e-12;
Matches 239; Conservative 0; Mismatches 246; Indels 6; Gaps 1;

QY 416 GGTCCCGAAGCTGGAATATGGAGCCAGACATGCGATGACGACTGCTTCTGTGCACTG 475
Db 869 GGGCCCGTACTGTGTGGCTTGGGCCACGCGTACACCCCGAGGAGTTGTGTGTAGCC 928
QY 476 GCTGTGAACAGCACTGGGCTCCCGTCTTTTGTGCCGACAAAGGCTGCTACTACTGCG 535
Db 929 AGTGTGGAGGTCCTGGAAGGGTGGCTTCTTTGAGGAGAAGGCGCCATCTTCTGCC 988
QY 536 TGCCCTGTATGAGAACAAAGTTTGTCTCTCGCTGGCGCGCTGCGAGCAAGCGCTGACAC 595
Db 989 CACCATGTATGACGTGGCTATGCACCCAGCTGTGCAAGTGCACCAAGAAGATTACAG 1048
QY 596 AGGTGGAGTGCATACCTGATCAGCGTGCATCGAGATGCTGTCTCTACCGGAT 655
Db 1049 -GCGAGATCATGACGCCCTGAAGATGACCTGGCAGCTGCACCTTTACCTGTGCTCT 1108
QY 656 GCCAGACGCCCTGGCAGGCGACAGTTTCACCTCCCGGATGAAGATCCCTACTGTGTGG 715
Db 1109 GCAAGACCCCATCCGAACAGGCGCTTCTACATGGAGGAGGCGCTGCTTATTCGGAGC 1168
QY 716 CCGTGTGGAGAACTCTTTGGACCTAAGTGCAGCAGCTGCAAGCGCCCATCGTAGGAC 775
Db 1169 GAGACTATGAGAAGATGTTGGCAGCAATGCTGCTGCTTCAAGATCG----- 1222
QY 776 TCGGTGGAGCAAGTATGTGCTTCTTGAAGCCGACACTGGCACCACCAACTCTCTCTCT 835
Db 1223 ACGTGGGAGCGCTTCTCTGGAGGCCCTTCGTGAGGCGCTTCTAGCTGGCATGACACCTGCTCTCT 1282

Qy	836	GCGCCGCTGCTTACTTCCTCGTGCGGACAGGGCTTGTACCGATGGAGACCAAGTGC	895
Db	1283	GTCGATATGTCAGATCAACCTGGAAGAAAGACCTTCTACTCCAAGAAGGACAGGCCTC	1342
Qy	896	TCTGCCAGGCG	906
Db	1343	TCTGCAAGAGC	1353

RESULT 12

```

US-09-124-238A-33
; Sequence 33, Application US/09124238A
; Patent No. 6300127
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6300127el Bone Mineralization Proteins, DNA, Vectors,
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/124,238A
; CURRENT FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 33
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-124-238A-33

```

```

RESULT 13
US-09-721-975-33
; Sequence 33, Application US/09721975
; Patent No. 6444803
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6444803el Bone Mineralization Proteins, DNA, Vectors,
; TITLE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/721,975
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/124,238
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 33
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-721-975-33

```

RESULT 14
US-09-986-621-33
; Sequence 33, Application US/09986621
; Patent No. 6521750
; GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 13:24:15 ; Search time 375 seconds
(without alignments)
12058.582 Million cell updates/sec

Title: US-09-786-135-2
Perfect score: 1971
Sequence: 1 ggtccctctgcagtcgacg.....ctcttgacgtggggatgaa 1971

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

*Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	429	21.8	431	11	US-09-918-995-36070
2	344.4	17.5	346	14	US-10-066-543-2286
3	252.2	12.8	2254	10	US-09-954-456-2244
4	244.4	12.4	2310	10	US-09-765-231A-31
5	232.6	11.8	830	11	US-09-764-891-1354
6	225	11.4	2154	13	US-10-044-090-260
7	152.8	7.8	580	10	US-09-833-381-786
8	117.6	6.0	445	11	US-09-918-995-4641
9	115.8	5.9	426	11	US-09-918-995-5309
10	112.4	5.7	424	11	US-09-918-995-4146
11	107.8	5.5	604	10	US-09-833-381-1725
12	99.8	5.1	724	10	US-09-833-381-1726
13	88.6	4.5	550	11	US-09-991-936-78
14	75.6	3.8	1486	9	US-09-854-864-3
15	75.4	3.8	1488	11	US-09-986-625-8
16	75.4	3.8	1620	11	US-09-986-625-32

Sequence 9, Appl	11	1644	3.8	17	75.4	3.8	1644	11	US-09-986-625-9
Sequence 33, Appl	11	1665	3.8	18	75.4	3.8	1665	11	US-09-986-625-33
Sequence 22, Appl	11	1689	3.8	19	75.4	3.8	1689	11	US-09-986-625-22
Sequence 32, Appl	9	1835	3.8	20	75.4	3.8	1835	9	US-09-925-302-32
Sequence 44, Appl	14	802	3.8	21	75	3.8	802	14	US-10-043-487-44
Sequence 28, Appl	14	3287	3.8	22	75	3.8	3287	14	US-10-175-523-28
Sequence 223, App	14	3287	3.8	23	75	3.8	3287	14	US-10-205-823-223
Sequence 221, App	14	6079	3.8	24	75	3.8	6079	14	US-10-205-823-221
Sequence 76, Appl	11	451	3.7	25	73	3.7	451	11	US-09-991-936-76
Sequence 23603, A	11	448	3.6	26	71.6	3.6	448	11	US-09-918-995-23603
Sequence 37, Appl	9	3502	3.5	27	69.4	3.5	3502	9	US-09-880-192-37
Sequence 7733, Ap	11	12103	3.4	28	67.8	3.4	12103	11	US-09-764-891-7733
Sequence 19419, A	9	303	3.4	29	67.4	3.4	303	9	US-09-864-761-19419
Sequence 8148, Ap	11	405	3.4	30	67.4	3.4	405	11	US-09-918-995-8148
Sequence 2701, Ap	9	476	3.4	31	67.4	3.4	476	9	US-09-864-761-2701
Sequence 3390, Ap	9	476	3.4	32	67.4	3.4	476	9	US-09-864-761-3390
Sequence 4613, Ap	10	480	3.4	33	67	3.4	480	10	US-09-878-574-4613
Sequence 2, Appl	11	1696	3.4	34	66.8	3.4	1696	11	US-09-986-625-2
Sequence 1, Appl	14	2060	3.4	35	66.6	3.4	2060	14	US-10-195-101-1
Sequence 1739, Ap	14	852	3.4	36	66.4	3.4	852	14	US-10-198-846-1739
Sequence 4381, Ap	10	498	3.3	37	65.8	3.3	498	10	US-09-878-574-4879
Sequence 4829, Ap	10	468	3.3	38	65.4	3.3	468	10	US-09-878-574-4381
Sequence 4177, Ap	10	468	3.3	39	65.2	3.3	468	10	US-09-878-574-4829
Sequence 39, Appl	13	481	3.3	40	65.2	3.3	481	13	US-09-878-574-4177
Sequence 4885, Ap	10	5700	3.3	41	65.2	3.3	5700	10	US-10-044-090-39
Sequence 4744, Ap	10	483	3.3	42	64.6	3.3	483	10	US-09-878-574-4885
Sequence 8634, Ap	14	463	3.3	43	64.4	3.3	463	14	US-09-878-574-4744
Sequence 11000, A	11	911	3.3	44	64.4	3.3	911	11	US-10-198-846-8634
	14	1102	3.3	45	64.4	3.3	1102	14	US-10-198-846-11000

ALIGNMENTS

RESULT 1

US-09-918-995-36070
; Sequence 36070, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36070
; LENGTH: 431
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(431)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-36070

Query Match	21.8%	Score 429;	DB 11;	Length 431;
Best Local Similarity	99.8%	Pred: No. 3.2e-124;		
Matches 429;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	424	AAGCTGGAATGATGAGGCCAGACATGCGCATGAGCAGCTGCTTCTGTGTCAGTGGCTGTGAA	483	
Db	1	AAGCTGGAATGATGAGGCCAGACATGCGCATGAGCAGCTGCTTCTGTGTCAGTGGCTGTGAA	60	
Qy	484	CAGCCACTGGGCTCCCGTCTTTTGTGCCGACAAAGGTCGTCTACTACTGTCGTCCTGTC	543	
Db	61	CAGCCACTGGGCTCCCGTCTTTTGTGCCGACAAAGGTCGTCTACTACTGTCGTCCTGTC	120	
Qy	544	TATGAGAACAAAGTTTGTCTCTCGTGCGCCGCTGTCAGCAGAGCTGACACAGGCTGGA	603	

Db 121 TATGAGAACAAAGTTTCTCTCTGCTGCGCCCGCTGCAGCAAGACGCTGACACAGGGTGA 180
QY 604 GTGACATACCGGTGATCAGCGGTGGCATCGAGAATGCTGGTCTGTACCGGATGCCAGACG 663
Db 181 GTGACATACCGGTGATCAGCGGTGGCATCGAGAATGCTGGTCTGTACCGGATGCCAGACG 240
QY 664 CCCCTGGCAGGCGACAGTTCACCTCCCGGGATGAAGATCCCTACTGTGTGGCCCTGTTTT 723
Db 241 CCCCTGGCAGGCGACAGTTCACCTCCCGGGATGAAGATCCCTACTGTGTGGCCCTGTTTT 300
QY 724 GGAGAACCTTTTGGACCTTAAGTGCAGCAGCTGCAGCGCCCGCCATCGTAGGACTCGGTGA 783
Db 301 GGAGAACCTTTTGGACCTTAAGTGCAGCAGCTGCAGCGCCCGCCATCGTAGGACTCGGTGA 360
QY 784 GCAAGTAGTGTCTCTTTGAAGACCGACACTGGCACCACAACTGCTTCTCTGCGCCCGC 843
Db 361 GCAAGTAGTGTCTCTTTGAAGACCGACACTGGCACCACAACTGCTTCTCTGCGCCCGC 420
QY 844 TGCTCTACT 853
Db 421 TGCTCTACT 430

RESULT 2
US-10-066-543-2286/c
; Sequence 2286, Application US/10066543
; Publication No. US2003008781A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuxiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2286
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-2286

Query Match 17.5%; Score 344.4; DB 14; Length 346;
Best Local Similarity 99.7%; Pred. No. 1.1e-97;
Matches 345; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1084 ACTCCTGACCCAGGCCCCCAATCTCTGGGCTTTACAGAGCTCCATGAGTCAAGCCGCC 1143
Db 346 ACTCCTGACCCAGGCCCCCAATCTCTGGGCTTTACAGAGCTCCATGAGTCAAGCCGCC 287
QY 1144 TCCACACACTGGAGTCCAGAAATTCACCTCTCCCTCGAGTCTGGGTCCAGACTGAG 1203
Db 286 TCCACACACTGGAGTCCAGAAATTCACCTCTCCCTCGAGTCTGGGTCCAGACTGAG 227
QY 1204 TCCTCTCCCAATCAGGGCTTAGACCCGAGCCCTCCAAACCTGGACTCTGGGACTTAG 1263
Db 226 TCCTCTCCCAATCAGGGCTTAGACCCGAGCCCTCCAAACCTGGACTCTGGGACTTAG 167
QY 1264 GCCCCCTTAATCTAGACTTCTCTTATAGTTTTCAGGTCTCCTATGGGTCCCTGGGAAG 1323
Db 166 GCCCCCTTAATCTAGACTTCTCTTATAGTTTTCAGGTCTCCTATGGGTCCCTGGGAAG 107
QY 1324 TCCTTGAAGTGGAGTGTCTCAGGCTTGACCTGCGCCCAACCCCATCCCGCGGTGAGGC 1383

Db 106 TCCTTGAAAGTGGAGTGTCTCAGGCTTGACCTGCCACCCATCCCGGGTTGAGGC 47
QY 1384 TGTGGGGCAGCAGATCAGGAGCCCACTGATAAGGGGCCCTAGGCT 1429
Db 46 TGTGGGGCAGTAGATCAGGAGCCCACTGATAAGGGGCCCTAGGCT 1

RESULT 3
US-09-954-456-2244
; Sequence 2244, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2244
; LENGTH: 2254
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-2244

Query Match 12.8%; Score 252.2; DB 10; Length 2254;
Best Local Similarity 57.1%; Pred. No. 2e-68;
Matches 480; Conservative 0; Mismatches 358; Indels 3; Gaps 1;

QY 77 TTGCCACCATGAGCGAGTCAATTTGACTGTGCAAAATGCAACGAGTCCCTGTATGACGCA 136
Db 76 TGGCGACCATGCGGAGAGTTTGAATGCCACTACTGCGAGGATCCCTTGCAGGGGAAGA 135
QY 137 AGTACATCCAGACAGACAGCGGCCCTACTGTGTCCCTGTGCTATGACAAATACCTTTGCCA 196
Db 136 AGTATGTCAAAAGAGTGGCCACCACTGTGCTGAAATGCTTTGACAAAGTTCTGTGCA 195
QY 197 ACACCTGTGCTGAGTGGCCAGCAGCTTATCGGCGCATGACTCGAGGGAGCTGTCTATGAG 256
Db 196 ACACCTGTGGAATGCGCGCAAGCCCATCGGTGGCGACTTCCAAAGAGGTGCACTATAAGA 255
QY 257 ACCGCAATTTCCACGAGGGCTGTTCCTGCTGTGCTGCTGCCAGGGCTCACTAGCCGATG 316
Db 256 ACCGCTTCTGCATGACACCTGCTTCCCTGTGCGCAAGTGCCTTCAACCCCTGGCCAATG 315
QY 317 AACCTTCCACCTGCCAGGACAGTGAAGTGCCTGTGCAATGACTGTCTACTGCACTGCGGT 376
Db 316 AGACCTTTGTGCCAAGGACCAACAGATCCTGTGCAACAAGTGCACCACTCGGGAGGACT 375
QY 377 CCTCGAGTGTCCGCTTGTGGGGAGACTGTCTATGCTGCTGGGTCCCGGAGCTGGAAATG 436

Db 376 TCCCAAGTGAAGGGGTGCTTCAAGGCCATTTGGCAGGAGATCAAAACGTGGAGTACA 435
Qy 437 GAGCCAGACATGACATGACACTGCTTCTCTGTCAGTGGTGTGAACAGCACTGGGCT 496
Db 436 AGGGAGCGTCTGGCAAGAGACTGCTTCACTAGTACTGCAAGCAAGTCAATCGGA 495
Qy 497 CCGTTCCTTTTGGCCCGCAAGAGGTGCTCACTACTGCGTCCCTGCTATGAGAACAGT 556
Db 496 CTGGAAGCTTCTTCCCTTAAGGGGAGGACTTCTACTGCTGACTTGGCATGAGACCAAGT 555
Qy 557 TTGCTCTCTGCTGGCCCGCTGCAGCAAGAGCTGACACAGGGTGGAGTGCATACACGTG 616
Db 556 TGGCAAGCATTTGGTGAAGTGCACAAAGCCATCAATCTGGAGGAATCACTTACCAGG 615
Qy 617 ATCAGCCGTGGCATGAGAAATGCTGTGCTGTACCCGGATGCCAGACGCCCTTGGCAGGC 676
Db 616 ATCAGCCCTGGCATGCGATTTGCTTGTGTGTACTGCTCTAAGAAGTGGCTGGC 675
Qy 677 AGCATTCACCTCCCGGGATGAAGATCCCTACTGTGTGGCTGTTTTGGAGAACTCTTTG 736
Db 676 AGCGTTTACCGCTGTGGAGGACCAGTATTACTGCGTGGATTGCTACAAGAACTTCGTGG 735
Qy 737 CACCTAAGTGCAGAGCTCAAGCGCCCATCTAGTACTGCTGTGGTGGAGG---CAAGTATG 793
Db 736 CCAAGAAGTGTGTGGATGCAAGAACCCCATCACTGGGTTTGGTAAAGGCTCCAGTGTGG 795
Qy 794 TGTCTTTTGAAGCCGACACTGGCACCACCAACTGCTTCTCTGCTGGCCCGCTGTCTACCT 853
Db 796 TGGCTATGAAGACATCTGGCAGCACTACTGCTTCCACTGCAAAAATGCTCCGTGA 855
Qy 854 CCGTGTGGCCAGGGCTTCCGTACCCGGATGGAGACCAAGTGTCTGTCCAGGGCTGTAGCC 913
Db 856 ATCTGGCCAAAGCGCTTTGTTTCCACCAGGAGCAAGTGTATTGTCCCGACTGTGCCA 915
Qy 914 A 914
Db 916 A 916

RESULT 4

US-09-765-231A-31
; Sequence 31, Application US/09765231A
; Patent No. US20020119452A1
; GENERAL INFORMATION:
; APPLICANT: Searle/Monsanto
; APPLICANT: Phippard, Deborah
; APPLICANT: Vasanthakumari, Geetha
; APPLICANT: Dotson, Stanton
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,
; FILE REFERENCE: SO-3221 PR
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 82
; SEQ ID NO 31
; LENGTH: 2310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-231A-31

Query Match 12.4%; Score 244.4; DB 10; Length 2310;
Best Local Similarity 57.2%; Pred. No. 5.7e-66;
Matches 482; Conservative 0; Mismatches 356; Indels 4; Gaps 2;
Qy 77 TTGCAACCATGAGGAGTATTTGACGTGCAAAATGCAACAGTCCCTGTATGGAGCA 136
Db 119 TGGCACCATTGGCGAGGAAGTTTACTGCTCCACTACTGAGGATCCCTTGCAGGGAGA 178
Qy 137 AGTACATCCAGACAGACAGCGCCCTTACTGTGTGCGCTGTATGACAACTTTGCCA 196
Db 179 AGTATGTCAAAAGGATGGCCACCACCTGCTGCCTGAAATGCTTGAAGTCTGTGCCA 238

Qy 197 ACACCTGTGTGAGTGGCAGCAGCTTATCGGGCATGACTCGAGGAGAGTGTCTTATGAG 256
Db 239 ACACCTGTGTGAGTGGCAGCAGCTTATCGGGCATGACTCGAGGAGAGTGTCTTATGAG 298
Qy 257 ACCGCCATTTCCAGGAGGCTGCTTCCGCTGCTCCGCTGCCAGCGTCACT-AGCCGAT 315
Db 299 ACCGCTTCTGGCATGACACCTGCTTCCGCTGTGCAAGTGTCTTACCCCTTGGCCCAAT 358
Qy 316 GAACCTTTACCTGCCAGGAGTGAAGTGTCTGTCAATGACTGCTGTCAGTGCCTTT 375
Db 359 GAGACCTTTTGGCCCAAGGACAAAGATCTGTGCAACAGTGCACACTCGGGAGGAC 418
Qy 376 TCTTCGAGTGTCTCCGCTTGTGGGAGAGTGTCTATGCTTGGGTCCCGGAAGCTGGAATAT 435
Db 419 TCCCCAAGTGAAGGGTGTCTTCAAGGCCATTTGTGCGAGGATCAAAAGTGGAGTAC 478
Qy 436 GGAGCCAGACATGGCATGAGCACTGCTTCTGTGAGTGGCTGTGAACGCCACTGGGC 495
Db 479 AAGGGGACCGTCTTGGCACAAGACTGCTTACCTGTAGTACTGCAAGCAAGTCAATCGG 538
Qy 496 TCCGCTTCTTTGTGCCCCGACAAAGGTGCTCACTACTGCGTCCCTGCTATGAGAACAA 555
Db 539 ACTGGAAGCTTCTTCCCTAAAGGGGAGGACTTCTACTGCGTACTTGCATGAGACCAAG 598
Qy 556 TTTGCTCTGCTGCTGCGCCCGCTGCAGCAAGACGCTGACACAGGGTGGAGTGACATACCGT 615
Db 599 TTTGCCAAGCATTCGCTGAAGTGCACAAAGGCCATCACATCTGGAGGAATCACTTACCAG 658
Qy 616 GATCAGCGTGGCATGAGAAATGCTGTGTGTACCGGATGCGGATGCCAGCCCTGGCAGGG 675
Db 659 GATCAGCGTGGCATGCGGATGCTTGTGTGTGTACTGCTCTAAGAAAGTGGCTGGG 718
Qy 676 CAGCAGTTCACCTCCCGGATGAAGATCCCTACTGCTGTGGCTGCTTTTGGAGAACTCTTT 735
Db 719 CAGGTTTCCCGCTGTGGAGGACCAATTTACTGCGTGGATTTCTACAAGAACTTTGTG 778
Qy 736 GCACCTAAGTGCAGCAGCTGCAAGCGCCCATCTAGGACTCTGGTGGAGG---CAAATAT 792
Db 779 GCCAAGAGTGTGCTGGATGCAAGAACCCCATCACTGGTGTGTAAGGCTCCAGTGTG 838
Qy 793 GTGTCTTTTGAAGACCGACACTGGCACCACACACTGCTTCTCTCGCCCGCTGCTCTACC 852
Db 839 GTGGCTTATGAAGGACAAATCCTGGCAGGACTACTGCTTCCACTGCAAAAATGCTCCGTG 898
Qy 853 TCCCTGTGGCCAGGCTTGTGATCCCGGATGGAGCAAGAGTGTCTGCGAGGCTGTAGC 912
Db 899 AATCTGCCCAACAGCGCTTTGTTTCCACAGGAGCAAGTGTATTGTCCGACTGTGCC 958
Qy 913 CA 914
Db 959 AA 960

RESULT 5

US-09-764-891-1354
; Sequence 1354, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1354
; LENGTH: 830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-1354
Query Match 11.8%; Score 232.6; DB 11; Length 830;

Best Local Similarity 57.5%; Pred. No. 2e-62;
Matches 409; Conservative 4; Mismatches 298; Indels 0; Gaps 0;

Qy		202	TGTCGTAGTGCCAGCAGCTTATACGGGCATATGACTCGAGGGAGCTGTCTTATGAAGACGCCG	261
Db		4	TGGCAGGAATGCAAAAAACCAATGAATCTGATTCTAAGGATCTTTGTTACAAGAAGACCGG	63
Qy		262	CATTTCCACAGGGCTGCTTCGGCTGCTCGCGCTGCCAGCGCTCACTAGCCGATGAACCC	321
Db		64	CATCGGCATGAAGGATCTTCAAAGTGACCANAATGCAATCACTCTTTGGTGGAAAGCCCT	123
Qy		322	TTCACTTGCAGGACAGCTGAGCTCTCTGCAATGACTGCTACTCGAGTGCCTTTTTCTCTCG	381
Db		124	TTTGCTGCCAAGNATGATGCGCGCTGCTGTGCACGGAGTGCTATTCTTAATGAGTGTCTCTCC	183
Qy		382	CAGTGCTCCGCTTGTGGGGAGACTGTCATGCCCTGGGTCGCCGAAGCTGGAATATGAGGC	441
Db		184	AAGTGCTTCCACTGCAAGAGAGACCATCATGCTGGTCCC GCCAAAATGGAATTTAAGGA	243
Qy		442	CAGNATGGCATGAGCACTGTTTCTCTGTGAGTGGCTGTGAAACAGCACACTGGGGCTCCGGT	501
Db		244	AAC TACTTGGCATGAWACCTGTTTGTGTGTGAGAAATTCGCCACAACTATAGSGACAAG	303
Qy		502	TCCTTTTGTGCCGACRAAGGTTGCTCACTACTTCGCTGCCCTGCTATGAGAACAACTTTGCT	561
Db		304	CTTTGATCTCCAAGAGATGGCAATATTGTGTGCCATGTTTKTGAGAAGGAGTTTGTCT	363
Qy		562	CCTCGCTGCCCGCTCCAGCAAGACGCTGACACAGGTTGGAGTGACATACCGTGATCAG	621
Db		364	CACTACTGCAACTTTTGTAGAAGTGATAACTTCAGTGGGATAACAATTTGTGACCAG	423
Qy		622	CCGTGGCATCGAATCTTGCTGTGTACCGGATGCCAGACGCCCTTGGCAGGGCAGCAG	681
Db		424	CTATGGCATAAAGAGTCKTTTCTGTGTAGTGGCTGTAGGAAAGATCTCTGTGAAGAACAG	483
Qy		682	TTCACTCCCGGATGAAGATCCCTACTGTGTGSCCTGTTTGGAGAACCTTTTGCACCT	741
Db		484	TTCACTGCCAGACGACTATCCATCTGCTGGACTGCTACAACTCTTTATGCGCAAC	543
Qy		742	AAGTGACAGCTGCAAGCGCCCCATCGTAGGACTCGGTGGAGGCAAGTATGTCTCTTT	801
Db		544	AAGTGTAGCCCTGTTCCAACCCATTTAGTGGTCTCACGGTGCCAAGTTTATCTGCTTT	603
Qy		802	GAAAGCCAGACTGGCACCAACAATGCTTCTCTGCGCCGCTGCTCTACCTCCCTGGTG	861
Db		604	CAAGACAGCCAGTGGCATAGGAATGCTTTAACTGCGGGAATGCTCTGTCTCTCTGGTG	663
Qy		862	GGCCAGGGCTTGTACCGGATGGAGACCAAGTGTCTTGCCAGGCTGTAGC	912
Db		664	GGTAAAGGCTTCTGTACCCAGAACAGGAATCTCTGCGCAAAAATGTGGG	714

RESULT. T 6

```

US-10-044-090-260
; Sequence 260, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 260
; LENGTH: 2154
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 361583.16
; NAME/KEY: unsure
; LOCATION: 2106.. 2150

```

US-10-044-090-260

	Query Match	11.4%;	Score 225;	DB 13;	Length 2154;
	Best Local Similarity	57.7%;	Pred. No. 6.8e-60;	Mismatches 295;	Indels 0; Gaps 0;
	Matches 402;	Conservative 0;	Mismatches 295;	Indels 0;	Gaps 0;
Qy	77	TTGCCACCATGAGCGAGTCATTGGACTGTGCNAATGCNACAGTCCCTGTATGACGCA	136		
Dd					
Dd	172	TGGGCACCATGGGAGGAAGTTTGACTGCCACTACTGCGAGGATCCCCTGCRAGGGAAGA	231		
Qy	137	AGTACATCCAGACAGACAGCGGCCCTACTGTGTGCCCTGCTATGACAATACTTTGCCA	196		
Dd					
Dd	232	AGTATGCAAAAGGATGCCACCACCTGCTCCTGAATGCTTGCAAGTTCTGTGCCA	291		
Qy	197	ACACCTGTGCTAGTGCCAGCAGCTTTATCGGSCATGATCTGAGGGAGCTGTGTTCTATGAAG	256		
Dd					
Dd	292	ACACCTGTGTGAATGGCGCAAGCCCATCGSTGCGGACTCCCAAGGAGGTGCATTAAGA	351		
Qy	257	ACCGCCATTTCACGAGGGCTGCTCCGCTGCTGCCGCTGCCAGCGCTCACATAGCCGATG	316		
Dd					
Dd	352	ACCGCTTCTGGCATGACACCTGCTTCCGCTGTGCCAAGTGCCCTCACCCCTTTGGCCAA	411		
Qy	317	AACCTCTTACCCTGGCCAGGACAGTGAGCTGCTCTGCAATGACTGCTACTGCATGCGCTTT	376		
Dd					
Dd	412	AGACCTTTGTGCCCAAGGACAACAAGATCCTGTGCACAAGTGCCACCCTCGGAGGACT	471		
Qy	377	CCTCGCAGTGTCCGCTTTGGGGAGACTGTGTCATGCCCTGGGTTCCGGGAAGCTGGAA	436		
Dd					
Dd	472	CCCCCAAGTGCAAGGGTGTCTTCAAGGGCCATTGTGGCAGGAGATCAAAACGTGGAGTACA	531		
Qy	437	GAGGCCAGACATGTCATCAGCAGCTCTTCCCTGTGCAGTGGCTGTGAACAGCCACATGGGCT	496		
Dd					
Dd	532	AGGGGACCGTCTGGCACAAGACCTGCTTACCTGTAGTAACGTGCAAGCAAGTCTATCGGGA	591		
Qy	497	CCCCTTCTTTTGCCCCGACAAAGGTGTCTCACTACTGGTGCCCTGCTATGAGAACAACT	556		
Dd					
Dd	592	CTGGAAGCTTCTCCCTAAAGGGGAGGACTTCTACTGCTGACTTTGCCATGAGACCAAGT	651		
Qy	557	TTGCTCTCGCTGGCCCGCTGCGACACAGACCTCACACAGGGTGGAGTGACATACCCGTG	616		
Dd					
Dd	652	TTGCGAAGCATPVTGGTGAAGTGCACAAGGCCATCACATCTGGAGGAATCACTTACCAGG	711		
Qy	617	ATCAGCCGTGGCATCGAAATGCTGTGTCGTGTACCGGATGCCAGCGCCCTGGCAGGCG	676		
Dd					
Dd	712	ATCAGCCCTGGCATGCCGATTGCTTTGTGTGTGTGTACCTGCTCTAAGAAGCTGGCTGGGC	771		
Qy	677	AGCAGTTTCACTCCCGGATGAAGATCCCTACTGTGTGGCCCTGTTTTGSAGAACTCTTTG	736		
Dd					
Dd	772	AGCGTTTCAACCGCTGTGGAGGACAGTATTACTGCGTGGATTGCTACAAGAATCTTTGTGC	831		
Qy	737	CACCTAAGTGACGAGCTGCAAGCGCCCCCATCGTAGG	773		
Dd					
Dd	832	CCAAGAAGTGTGCTGGATGCAAGAACCCCACTACATGG	868		

RESULT 7

```

US-09-833-381-786/c
; Sequence 786, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 786
; LENGTH: 580
; TYPE: DNA

```

; ORGANISM: Homo sapiens
US-09-833-381-786

Query Match: 7.8%; Score 152.8; DB 10; Length 580;

Best Local Similarity 56.9%; Pred. No. 2e-37; Mismatches 212; Indels 0; Gaps 0;

Matches 280; Conservative 0;

```
QY 421 CGGAAGCTGGAATATGAGGCCAGACATGGCATGGCACTGCTTCCTGTGCACTGGCTGT 480
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 580 CGCAAAATGGAATTTAAGGAACTACTGGCATGATACCTGTTCTGTGTGAGATGTC 521
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 GACAGCCACTGGGCTCCGCTTTTGTGCGCCGACGAAGGCTGCTCACTACTCGGTCGCC 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 520 CGACAACTTATACGCAAAAGCCTTTGATCTCCAAAAGAGAGTGGCAATATTGTGTCGA 461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 TGTATGAGAACAAGTTTGTCTCTGCTGCGCCGCTGCAGCAAGAGCTGACACAGGT 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 460 TGTGTGAGAGAGTTTGTCTCACTACTGCACTTTTGTAGAGGTGATACTTCAGGT. 401
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 GGAGTGACATPACGTCATCAGCGTGGCATCGAGAAATGCTGTCTGTACCGGATGCCAG 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 400 GGGATAACATTTTGTGACCACTATGGCATAAAGAGTGTCTGTGTAGTGTGCTGTAGG 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 ACSCCCTGGCAGGCGAGTTCACTCCCGGGATGAGATCCCTACTGTGTGGCTGT 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 340 AAAGATCTCTGTGAAGAACAAGTTCAATGTCAGAGAGCACTATCCATTCTGCATGGACTGC 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 TTTGGAGAATCTTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATCGTAGCACTCGGT 780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 280 TACAACCACTTTATGCCAACAAAGTGTAGCTGTTCCTCCAAACCCATTTAGTGTCTCACA 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 781 GGAGGCAAGTATGTCTTTGAAGACCGACACTGGCACCACCACTCTCTCTCGCGCC 840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220 GGTGCCAAGTTTATCTGCTTCAAGACAGCCAGTGGCATAGCGAATGCTTAACTGGGG 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 841 CGTGTCTACCTCCCTGTGTGGCCAGGCTTGTACCGGATGGAGCAAGTGTCTGTGC 900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 160 AAATGTCTGTCTCTCTGTGTGGTAAAGGCTTCTGACCCAGCAACAAGAAATCTTCTGC 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 901 CAGGCGCTGTAGC 912
    ||||| |||||
Db 100 CAAAATGTGGC 89
```

RESULT 8

US-09-918-995-4641
; Sequence 4641, Application US/09918995
; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 4641

; LENGTH: 445

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(445)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-4641

Query Match

Best Local Similarity 6.0%; Score 117.6; DB 11; Length 445;

Matches 201; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

```
QY 77 TTGCCACCATGAGCGAGTCAATTTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGCA 136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96 TGGCACCATTGGCGGAGAAGTTTGAATGCCACTACTGCAGGATCCCTTGCAGGGGAAGA 155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 137 AGTACATCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTATGACAATACCTTTGCCA 196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 156 AGTATGTCAAAAGGATGGCCACCACTGCTGCCTGAAATGCTTTGACAAGCTCTGTGCCA 215
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 197 ACACCTGTGTGAGTGGCCACAGCTTATCGGGCATGACTCGAGGAGCTGTTCTATGAAG 256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 216 ACACCTGTGTGGAATCGCGAAGCCCATCGTGGGACTCAAGAGGTGCACATATAGA 275
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 257 ACCGCCATTTCCACGAGGGTGTCTTCGCTGTCTGCCCTGCCAGCGTCACTAGCCGATG 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 276 ACCGCTTCTGGCATGACACCTGTCTCGCTGTGCAAGTGCCTTACCCCTTGGCCAAATG 335
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 317 AACCTTCACTGCCAGGACAGTGAAGTGTCTGCAATGACTGTCTACTGAGTGCCTTTT 376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 336 AGACCTTTGTGGCCAAAGGACACAAGATCTGTGCAACAAGTGCACCACTCGGAGGACT 395
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 377 CCTCGCAGTGTCTCGCTTTGTGGGAGACTGTCATGCCTGG 416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 396 CCCCCAAGTCAAGGGGTGCTTCAAGGCCAATTGTGGCAGG 435
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 9

US-09-918-995-5309

; Sequence 5309, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 5309

; LENGTH: 426

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(426)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-5309

Query Match

Best Local Similarity 5.9%; Score 115.8; DB 11; Length 426;

Matches 192; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

```
QY 77 TTGCCACCATGAGCGAGTCAATTTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGCA 136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96 TGGCACCATTGGCGGAGAAGTTTGAATGCCACTACTGCAGGATCCCTTGCAGGGGAAGA 155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 137 AGTACATCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTATGACAATACCTTTGCCA 196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 156 AGTATGTCAAAAGGATGGCCACCACTGCTGCCTGAAATGCTTTGACAAGTCTGTGCCA 215
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 197 ACACCTGTGTGAGTGGCCACAGCTTATCGGGCATGACTCGAGGAGCTGTTCTATGAAG 256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 216 ACACCTGTGTGGAATCGCGAAGCCCATCGTGGGACTCAAGAGGTGCACATATAGA 275
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 257 ACCGCCATTTCCACGAGGGTGTCTTCGCTGTCTGCCCTGCCAGCGTCACTAGCCGATG 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 276 ACCGCTTCTGGCATGACACCTGTCTCGCTGTGCAAGTGCCTTACCCCTTGGCCAAATG 335
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 317 AACCTTCACTGCCAGGACAGTGAAGTGTCTGCAATGACTGTCTACTGAGTGCCTTTT 376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 336 AGACCTTTGTGGCCAAAGGACACAAGATCTGTGCAACAAGTGCACCACTCGGAGGACT 395
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

; TYPE: DNA

217 CAGCTTATCGGGCATGACTCGAGGGAGCTGTTCTATGAGACCGCCATTTCACGAGGGC 276


```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-986-625-8

Query Match      3.8%; Score 75.4; DB 11; Length 1488;
Best Local Similarity 48.7%; Pred. No. 5.8e-13;
Matches 239; Conservative 0; Mismatches 246; Indels 6; Gaps 1;

QY 416 GGTCCCGGAAGCTGGAATATGGAGCCGACATGGCAGCACTGCTTCCTGTGCAGTG 475
   || || || || || || || || || || || || || || || || || || ||
Db 713 GGGGCGCTACTGTGTGGGTTGGCCACGCGTACACCCGGAGGAGTTGTGTAGCC 772
   || || || || || || || || || || || || || || || || || || ||

QY 476 GCTGTGAACAGCACTGGGCTCCCGTCTTTTGTGCCGACAAAGGTGCTCACTACTGCG 535
   || || || || || || || || || || || || || || || || || || ||
Db 773 AGTGTGGGAAGTCTCTGGAAGAGGTGGCTTCTTTGAGGAGAAGGGCGCATCTTCTGCC 832
   || || || || || || || || || || || || || || || || || || ||

QY 536 TCCCTTGCTATGAGAACAAAGTTTGTCTCTCGCTGGCGCGCTGCAGCAAGAGCGCTGACAC 595
   || || || || || || || || || || || || || || || || || || ||
Db 833 CACCATGTATGACGTGCGCTATGACCCAGCTGTGCCAAGTGCAAGAAGATTACAG 892
   || || || || || || || || || || || || || || || || || || ||

QY 596 AGGTTGGAGTACATACCGTGATCAGCGTGGCATCGAGATGCTGTCTGTACCGGAT 655
   || || || || || || || || || || || || || || || || || || ||
Db 893 GCGAGATCATGCACGCCCTGAAGATGACCTGGCAGCTGCACCTGTACCTGTGTGCT 952
   || || || || || || || || || || || || || || || || || || ||

QY 656 GCCAGACGCCCTGGCAGGGCAGCAGTTCACTCCCGGGATGAAGATCCCTACTGTGTGG 715
   || || || || || || || || || || || || || || || || || || ||
Db 953 GCAAGACGCCCATCCGGAACAGGGCTTCTACATGGAGAGGGCGTGCCTATTGCGAGC 1012
   || || || || || || || || || || || || || || || || || || ||

QY 716 CCTGTTTGGAGAACTCTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATCGTAGGAC 775
   || || || || || || || || || || || || || || || || || || ||
Db 1013 GAGACTATGAGAAGATGTTGGCAGAAATGCCATGGCTGTGACTTCAAGATCG----- 1066
   || || || || || || || || || || || || || || || || || || ||

QY 776 TCGGTGGAGGCAAGTATGTCTCTTTGAAGACCGCACTGGCACCACAACTGCTTCTCT 835
   || || || || || || || || || || || || || || || || || || ||
Db 1067 AGGCTGGGACCGCTTCTCTGGAGGCCCTGGGCTTCAGCTGGCATGACACCTGCTTCGTCT 1126
   || || || || || || || || || || || || || || || || || || ||

QY 836 GGGCCCGTGTCTTACCTCCCTGGTGGCCAGGGCTTCGTACCCGATGGAGACCAAGTGC 895
   || || || || || || || || || || || || || || || || || || ||
Db 1127 GTGCGATATGTGAGATCAACCTGGAAGGAAGACCTTCTACTCCCAAGAGGACAGGCGCTC 1186
   || || || || || || || || || || || || || || || || || || ||

QY 896 TCTGCCAGGGC 906
   || || || || || ||
Db 1187 TCTCAAGAGC 1197
   || || || || || ||
```

Search completed: August 27, 2003, 15:32:24
Job time : 376 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 11:50:19 ; Search time 2837 Seconds

(without alignments)
16885.486 Million cell updates/sec

Title: US-09-786-135-2

Perfect score: 1971

Sequence: 1 gdtccgtctgcagtcgcag.....ctcttgacgtggggatgaa 1971

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_estl:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	988.2	50.1	1201	13	BX385643
2	970.8	49.3	1172	9	AL527609
3	957	48.6	1201	9	AL523732
4	951.6	48.3	1201	13	BX437971
					BX437971 BX437971

C	5	951.4	48.3	1082	9	AL526160
C	6	942.6	47.8	1201	9	AL528549
C	7	939.4	47.7	1201	9	AL525660
C	8	922.8	46.8	1200	13	BX345083
C	9	917.6	46.6	1201	9	AL526865
C	10	903	45.8	1201	9	AL523733
C	11	900.6	45.7	1200	9	AL526734
C	12	894.6	45.4	1201	13	BX385642
C	13	894.2	45.4	1116	9	AL551558
C	14	894.2	45.4	1201	13	BX445770
C	15	888.4	45.1	1034	12	BM480273
C	16	888.4	45.1	1044	12	BM557043
C	17	878.4	44.6	1201	9	AL563163
C	18	865.4	43.9	1096	9	AL550299
C	19	861.4	43.7	979	12	BM545780
C	20	836.8	42.5	1201	13	BX324699
C	21	830	42.1	877	9	AL521288
C	22	826.6	41.9	997	12	BI757781
C	23	816.4	41.4	1201	9	AL528415
C	24	812.8	41.2	1201	9	AL528520
C	25	811	41.1	933	9	AL521265
C	26	810.8	41.1	883	12	BM008985
C	27	810.6	41.1	1201	13	BX445771
C	28	807.2	41.0	853	12	BI822670
C	29	796.4	40.4	1201	9	AL528550
C	30	795.8	40.4	1071	9	AL550271
C	31	768.4	39.0	1043	9	AL525946
C	32	767.8	39.0	1125	9	AL562940
C	33	767.2	38.9	773	14	CA447438
C	34	765	38.8	1191	13	BX355333
C	35	764.2	38.8	855	12	BM010854
C	36	760.6	38.6	909	13	BQ675737
C	37	760.6	38.6	955	12	BG826380
C	38	756.8	38.4	1201	9	AL528387
C	39	756.2	38.4	1044	12	BM924990
C	40	747.6	37.9	1165	9	AL526900
C	41	746	37.8	865	12	BI908150
C	42	745	37.8	909	13	BX374168
C	43	744.6	37.8	1201	9	AL526613
C	44	741.8	37.6	891	13	BX372660
C	45	739.8	37.5	1145	9	AL551594

ALIGNMENTS

RESULT 1	BX385643	1201 bp	mrna	linear	EST 08-MAY-2003
LOCUS	BX385643	Homo sapiens	NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens		
DEFINITION	CDNA clone CS0DC014YA02 5-PRIME, mRNA sequence.				
ACCESSION	BX385643				
VERSION	BX385643.1	GI:30440553			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
-ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1201)				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
COMMENT	Contact: Genoscope				
	Genoscope - Centre National de Sequencage				
	BP 191 91006 EVRY cedex - France				
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr				
	Library was constructed by Life Technologies, a division of				
	Invitrogen. This sequence belongs to sequence cluster 888.f For				
	more information about this cluster, see				
	http://www.genoscope.cns.fr/				
	cgi-bin/cluster.cgi?seq=CSIAC0042C02QPI&cluster=888.f. Contact :				
	Feng Liang Email : fliang@lifetech.com URL :				
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600				


```
Db 130 GCA--GTTCTCTCCCTGTCACCATGAGCAGTCAATTTGACGTGTGCAAAATGCAACGAG 187
QY 121 TCCTGTATGAGCGCAAGTATACCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTAT 180
Db 188 TCCTGTATGAGCGCAAGTATACCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTAT 247
QY 181 GACAATACCTTTGCCAACACCTGTGTGAGTGCAGCAGCTTATCGGGCATGACTGAGG 240
Db 248 GACAATACCTTTGCCAACACCTGTGTGAGTGCAGCAGCTTATCGGGCATGACTGAGG 307
QY 241 GAGCTGTTTATGAAGACCGCCATTTCCAGAGGGCTGCTTCGCTGCTGCTGCCGTGCCAG 300
Db 308 GAGCTGTTTATGAAGACCGCCATTTCCAGAGGGCTGCTTCGCTGCTGCTGCCGTGCCAG 367
QY 301 CGCTCACTAGCCGATGAACCTTCACCTGCCAGGAGTGCAGTGCCTGCTCAATGACTGC 360
Db 368 CGCTCACTAGCCGATGAACCTTCACCTGCCAGGAGTGCAGTGCCTGCTCAATGACTGC 427
QY 361 TACTGAGTGCCTTTCTCGCAGTGTCTCGCTGCTGGGAGACTGTCTGCTGCTGCTGCC 420
Db 428 TACTGAGTGCCTTTCTCGCAGTGTCTCGCTGCTGGGAGACTGTCTGCTGCTGCTGCC 487
QY 421 CGGAAGCTGGAATATGAGGCGCCAGACATGCGATGAGCACTGCTTCCTGTGCTGCTGCT 480
Db 488 CGGAAGCTGGAATATGAGGCGCCAGACATGCGATGAGCACTGCTTCCTGTGCTGCTGCT 547
QY 481 GAACGCCACTGCGTCCCGTCTTTGTCGCCGACAGAGGTCCTCACTACTGCTGCCGTC 539
Db 548 GAACGCCACTGCGTCCCGTCTTTGTCGCCGACAGAGGTCCTCACTACTGCTGCCGTC 607
QY 540 CTGCTATGAGAACAAAGTTTCTCCTCGCTGCGCCGCTGCAGCAAGCAGTGCACAGGG 599
Db 608 CTGCTATGAGAACAAAGTTTCTCCTCGCTGCGCCGCTGCAGCAAGCAGTGCACAGGG 667
QY 600 TGGAGTGACATACCGTATGATCAGCGCTGGCATGAGCAATGCTGCTGTGTACCGGATGCCA 659
Db 668 TGGAGTGACATACCGTATGATCAGCGCTGGCATGAGCAATGCTGCTGTGTACCGGATGCCA 727
QY 660 GAGCGCCCTGGCAGGCGAGTTCACCTCCCGGATGAAGATCCCTACTGCTGCCGTCG 719
Db 728 GAGCGCCCTGGCAGGCGAGTTCACCTCCCGGATGAAGATCCCTACTGCTGCCGTCG 787
QY 720 TTTTGGAGACTCTTTGCACTAAGTGACAGCTGCAAGCGCCGCTGCTAGGACTCGG. 779
Db 788 TTTTGGAGACTCTTTGCACTAAGTGACAGCTGCAAGCGCCGCTGCTAGGACTCGG 847
QY 780 TGGAGGCAAGTATGTCTTTGAGACCGACACTGCGACCCCAACTGCTTCTCCTGCCG 839
Db 848 TGGAGGCAAGTATGTCTTTGAGACCGACACTGCGACCCCAACTGCTTCTCCTGCCG 907
QY 840 CCGCTGCTCTACCTCCTGCTGGGCGAGGCTTCGTACCGATGGAGACCAAGTCTGCTG 899
Db 908 CCGCTGCTCTACCTCCTGCTGGGCGAGGCTTCGTACCGATGGAGACCAAGTCTGCTG 967
QY 900 CCAGGGCTGAGCAGGCGCCCTAAGCAGGCGCTCTGAGCCAGGCTTTCCCATAC 959
Db 968 CCAGGGCTGAGCAGGCGCCCTAAGCAGGCGCTCTGAGCCAGGCTTTCCCATAC 1027
QY 960 CAGGGCGCCAGGACTGTGGCTCTTTTCTAAACCACTCTGAGGACTCAGCTCCCGCGCC 1019
Db 1028 ACGGGCGCCAGGACTGTGGCTCTTTTCTAAA--CACCTCTGGGACTCAGCTCCCGCGCC 1086
QY 1020 AAAAAAATGGGTCTCTCTGCTGGCTCCAGGATTTGCTCCCC 1061
Db 1087 AAAAAAATGGGTCTCTCTGCTGGCTCCAGGATTTGCTCCCC 1124
```

RESULT 3

AL523732/c

LOCUS

DEFINITION AL523732 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC003YC05 3-PRIME, mRNA sequence.

ACCESSION

AL523732.2 GI:31041993

VERSION

EST.

KEYWORDS

Homo sapiens (human)

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ORGANISM

1 (bases 1 to 1201)

REFERENCE

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

AUTHORS

Full-length cDNA libraries and normalization

JOURNAL

COMMENT

On Feb 13, 2001 this sequence version replaced gi:12787225.

Contact:

Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 888.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DC003AB03NP1&cluster=888.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/

Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DC003AB03NP1.

FEATURES

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DC003YC05"

/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 269 a 289 c 364 g 211 t 68 others

ORIGIN

Query Match

Best Local Similarity 48.68; Score 957; DB 9; Length 1201;

Matches 986; Conservative 32; Mismatches 23; Indels 4; Gaps 3;

QY 504 TTTTGTGCCCCGACAAAGGGTGTCTACTGTGGTCCCTGCTATGAGAACAAAGTTTGTCTCC 563

Db 1041 TTTTGTGCCCCGACAAAGGGTGTCTACTGTGGTCCCTGCTATGAGAACAAAGTTTGTCTCC 983

QY 564 TCGTGTGCCCCGCTGTCAGCAAGAGCTGTACACAGGGTGGAGTGACATACCGTGTATCAGCC 623

Db 982 CTCGGTGTCCGG-TGCAGCAARACGTTTACACAGGGTGGAGTGACATACCGTGTATCAGCC 924

QY 624 GTGGCATCGAAGATGTCTGTCTACCGGATGCCAGACGCCCTTGGCAGGGCAGCAGTT 683

Db 923 GTGGCATCGAAGATGTCTGTCTACCGGATGCCAGACGCCCTTGGCAGGGCAGCAGTT 864

QY 684 CACCTCCCGGATGAAGATCCCTACTGTGGTCCCTGTTTGGAGAACTCTTTGACACCTAA 743

Db 863 CACCTCCCGGATGAAGATCCCTACTGTGGTCCCTGTTTGGAGAACTCTTTGACACCTAA 804

QY 744 GTGACAGCAGTGCAGCGCCCATCTCGTAGGACTCGGTGGAGGCAAGTATGTCTCTTGA 803

Db 803 GTGCRGACGTGCAAGCGCCCATCTCGTAGGACTCGGTGGAGGCAAGTATGTCTCTTGA 744

QY 804 AGACCGACACTGGCACCACCAACTGCTTCTCTCGCGCCGCTGCTACTACCTCCCTGGTGGG 863

Db 743 AGACCGAGACTGGCACCACCAACTGCTTCTCTCGCGCCGCTGCTACTACCTCCCTGGTGGG 684

QY 864 CCAGGGCTTCTGATCCGGATGGAGACCAAGTGTCTGACAGGGCTGTAGCCAGGCGAGGCC 923

Db 683 CCAGGGCTTCTGATCCGGATGGAGACCAAGTGTCTGACAGGGCTGTAGCCAGGCGAGGCC 624

QY 924 CTAGCCAGGGCTCTCTGGACCCAGGCTTCCCATACACACGGGCCCGGAGGCTGTGGCTCT 983

```
Db 623 CTAAGCCAGGGCTCTGGACCCAGGCTTTCCATACACAGGGCCAGGACTGTGGCTCT 564
QY 984 TTTCTAAACCACTCTGGGACTCAGCTCCCCCGCCAAAATAATGGGTCTCTCTGGG 1043
Db 563 TTTCTAAACCACTCTGGGACTCAGCTCCCCCGCCAAAATAATGGGTCTCTCTGGG 504
QY 1044 CTCAGGATGTCTCCCACTCCAGATCCCAAACTGGTACTCCCTGACCCAGGCCGCC 1103
Db 503 CTCAGGATGTCTCCCACTCCAGATCCCAAACTGGTACTCCCTGACCCAGGCCGCC 444
QY 1104 AATCTGTGGCTTTACAGAGCTCCATGATGTAAGCCGCCCTCCCAACACTGGAGCTCAG 1163
Db 443 AATCTGTGG--TCTTACAGAGCTCCATGATGTAAGCCGCCCTCCCAACACTGGAGCTCAGG 386
QY 1164 AATTCACACCTCTCCCTCGAGTCTGGGTTCCAGACTGAGTCTCTCCCAAACTCAGGSC 1223
Db 385 KATTCACACCTCTCCCTCGAGTCTGGGTTCCCAACTGAGTCTCTCCCAAACTCAGGSC 326
QY 1224 TCTAGACCCGAGCCCTCCAAACCTGGACTCTGGGACTTTAGGCCCTTTAAATCTAGACTT 1283
Db 325 TCTAGACCCGAGCCCTCCAAACCTGGACTCTGGGACTTTAGGCCCTTTAAATCTAGACTT 266
QY 1284 CTCCTTATAGTTTCAGTCTCTCTATGAGTCTGGGAGTCTTGAAGTGGAGTCTTC 1343
Db 265 CTCCTTATAGTTTCAGTCTCTCTATGAGTCTGGGAGTCTTGAAGTGGAGTCTTC 206
QY 1344 TCAGGCTTGACCTGCCCAACCCCAATCCCGGGGTTGAGGCTGTGGGGCAGCAGATCAGG 1403
Db 205 TCAGGCTTGACCTGCCCAACCCCAATCCCGGGTGTGAGGCTGTGGGGCAGCAGATCAGG 146
QY 1404 AGCCACCTGATAAGGGGCCCTAGGTACAGGCTGTGCCAGCAGGTGCCACCCAGTGT 1463
Db 145 AGCCACCTGATAAGGGGCCCTAGGTACAGGCTGTGCCAGCAGGTGCCACCCAGTGT 86
QY 1464 CTCCTCATTTTATTTACAGTCTCATTTTCCCATAGATGGCAGAGGGTGAGATTGGCTC 1523
Db 85 CTCCTCATTTTATTTACAGTCTCATTTTCCCATAGATGGCAGAGGGTGAGATTGGCTC 26
QY 1524 ATCCCTCTCCAGATTCGCAATAA 1548
Db 25 ATCCCTCTCCAGATTCGCAATAA 1
```

RESULT 4

```
LOCUS BX437971
DEFINITION BX437971 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CS0DA002YG10 5-PRIME, mRNA sequence.
ACCESSION BX437971
VERSION BX437971.1 GI:30785652
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 888.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DA002BD05Qpl&cluster=888.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DA002BD05Qpl.
FEATURES Location/Qualifiers
1..1201 source
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA002YG10"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 236 a 373 c 315 g 257 t 20 others
ORIGIN
Query Match 48 3%; Score 951.6; DB 13; Length 1201;
Best Local Similarity 97.7%; Pred. No. 2.2e-231;
Matches 993; Conservative 3; Mismatches 14; Indels 6; Gaps 3;
QY 22 CTCGCCGAGCTCGCTCGGCCCGCCCGCTTGGCCCGCCCGAGGGTTCTCTCCCTTGGC 81
Db 64 CTCGCCGAGCTCGCTCGGCCCGCCCGCTTGGCCCGCCCGAGT-NNCTCTCCCTTGGC 122
QY 82 ACCATGAGGAGTCAATTTGACTGTGCAAAATGCAAGAGTCCCTGATGGACGCAAGTAC 141
Db 123 ACCATGAGGAGTCAATTTGACTGTGCAAAATGCAAGAGTCCCTGATGGACGCAAGTAC 182
QY 142 ATCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTATGACAATACCTTTGCCAACAC 201
Db 183 ATCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTATGACAATACCTTTGCCAACAC 242
QY 202 TGTGCTGAGTCCAGCAGCTTATCGGGCATGACTCGAGGAGCTGTTCTATGAAGACCGC 261
Db 243 TGTGCTGAGTCCAGCAGCTTATCGGGCATGACTCGAGGAGCTGTTCTATGAAGACCGC 302
QY 262 CATTTCCAGAGGGTCTCGCGCTGCCCTGCCAGCGCTCACTAGCCGATGACCC 321
Db 303 CATTTCCAGAGGGTCTCGCGCTGCCCTGCCAGCGCTCACTAGCCGATGACCC 362
QY 322 TTCACCTGCCAGGAGTGTGCTGCTGCAATGACTGTACTGAGTGTGCTTTTCTCG 381
Db 363 TTCACCTGCCAGGAGTGTGCTGCTGCAATGACTGTACTGAGTGTGCTTTTCTCG 422
QY 382 CAGTGTCCGCTTGTGGGAGACTGTCTATGCTGGTCCCGAAGCTTGAATATGAGGC 441
Db 423 CAGTGTCCGCTTGTGGGAGACTGTCTATGCTGGTCCCGAAGCTTGAATATGAGGC 482
QY 442 CAGACATGSCATGAGCAGTCTTCTGCTGCAAGTGTGTGAACACGCTGCGGCTCCCGT 501
Db 483 CAGACATGSCATGAGCAGTCTTCTGCTGCAAGTGTGTGAACACGCTGCGGCTCCCGT 542
QY 502 TCTTTTGTCCCGACAAAGGGTGTCTCAGTGTGCTGCTGCTATGAGAACAAGTTGCT 561
Db 543 TCTTTTGTCCCGACAAAGGGTGTCTCAGTGTGCTGCTGCTATGAGAACAAGTTGCT 602
QY 562 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
Db 603 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662
QY 622 CCGTGGCATCGAGATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 681
Db 663 CCGTGGCATCGAGATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 722
QY 682 TTCACCTCCCGGATGAAGATCCCTTACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
Db 723 TTCACCTCCCGGATGAAGATCCCTTACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
QY 742 AAGTGCAGCAGTGCAGCGGCCCATCTAGAGCTGCGTGGAGGCAAGTATGTGCTTTT 801
Db 783 AAGTGCAGCAGTGCAGCGGCCCATCTAGAGCTGCGTGGAGGCAAGTATGTGCTTTT 838
QY 802 GAAGACCCAGACTGCGACCAACAGTCTTCTCTGCGCCCGCTGCTGCTGCTGCTGCTGCTG 861
Db 839 GAAGACCCAGACTGCGACCAACAGTCTTCTCTGCGCCCGCTGCTGCTGCTGCTGCTGCTG 898
```

QY 862 GGCAGGCTTGTACCGATGGAGACCAAGTGTCTGCGCAGGCTGTAGCCAGGAGG 921
 D 899 GGCAGGCTTGTACCGATGGAGACCAAGTGTCTGCGCAGGCTGTAG-CAGGAGG 957
 QY 922 CCTAAGCCAGGCTCTGAGCCAGGCTTTCCATACCAAGGCGCCAGGACTGTGCTC 981
 D 958 CCCTAAGCCAGGCTCTGAGCCAGGCTTTCCATACCAAGGCGCCAGGACTGTGCTC 1017
 QY 982 CTTTCTAAACCACTCTGGACTCAGCTCCCGCCCAAAAAAATGGCTCTCCT 1037
 D 1018 CTTTCTAAACCACTCTGGACTCAGCTCCCGCCCAAAAAAATGGCTCTCCT 1073

RESULT 5
 AL526160/c
 LOCUS
 DEFINITION AL526160 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 CDNA clone CS0DC015YK07 3-PRIME, mRNA sequence.

ACCESSION AL526160
 VERSION AL526160.2 GI:31064021
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 TITLE 1 (bases 1 to 1082)
 JOURNAL Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 COMMENT Full-length cDNA libraries and normalization
 Unpublished
 On Feb 13, 2001 this sequence version replaced gi:12789653.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 888.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DC015AF04NP1&cluster=888.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DC015AF04NP1.
 Location/Qualifiers
 1. .1082

FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="CS0DC015YK07"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(4T)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 235 a 274 c 348 g 199 t 26 others

BASE COUNT
 ORIGIN
 Query Match 48.3%; Score 951.4; DB 9; Length 1082;
 Best Local Similarity 96.9%; Pred. No. 2.5e-231;
 Matches 988; Conservative 11; Mismatches 18; Indels 3; Gaps 3;
 QY 530 ACTCGTGCCTGTATGAGAACAAAGTTGCTCTCGCTGCGCCGCTGCAGCAAGACGC 589
 D 1017 ACTACGGTGCCTGCTWKRGRACAAAGTTGCTCC-CTCGGGCCCGCTGCAGCAAGACGC 959
 QY 590 TGACACAGGTGGAGTGACATACCGTGTATCAGCGGTGGCATCGGAATGTCTGTCTGTA 649
 D 958 TGACACAGGTGGAGTGACATACCGTGTATCAGCGGTGGCATCGGAATGTCTGTCTGTA 899
 QY 650 CCGGATCCAGACGCCCTCGCAGGCGGAGGAGTTCACCTCCCGGGATGAGATCCCTACT 709
 D 898 CCGGATCCAGACGCCCTCGCAGGCGGAGGAGTTCACCTCCCGGGATGAGATCCCTACT 839

RESULT 6
 AL528549/c
 LOCUS
 DEFINITION AL528549 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 CDNA clone CS0DD001YE04 3-PRIME, mRNA sequence.
 ACCESSION AL528549
 VERSION AL528549
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 TITLE 1 (bases 1 to 1201)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 FULL-LENGTH cDNA LIBRARIES AND NORMALIZATION

QY 710 GTGTGGCTCTTTTGGAGAACTCTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATCG 769
 D 838 GTGTGGCTCTTTTGGAGAACTCTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATCG 779
 QY 770 TAGGACTCGGTGGAGCAAGTATGTCTTTTGAAGACCGACACTGGCACCACAACTGCT 829
 D 778 TAGGACTCGGTGGAGCAAGTATGTCTTTTGAAGACCGACACTGGCACCACAACTGCT 719
 QY 830 TCTCTCGGCGCGCTCTACCTTCCCTGGTGGCCAGGCGCTTGTACCGGATGGAGACC 889
 D 718 TCTCTCGGCGCGCTCTACCTTCCCTGGTGGCCAGGCGCTTGTACCGGATGGAGACC 659
 QY 890 AAGTGCTCTGCCAGGCGGTAGCCAGGCGGCGCTTAAGCCAGGCGCTCTTGACCCAGGC 949
 D 658 AAGTGCTCTGCCAGGCGGTAGCCAGGCGGCGCTTAAGCCAGGCGCTCTTGACCCAGGC 599
 QY 950 TTTCCCATACACAGGCGCCAGGACTGTGGCTCTTTTCTAAACCACTCTGGGACTCAGC 1009
 D 598 TTTCCCATACACAGGCGCCAGGACTGTGGCTCTTTTCTAAACCACTCTGGGACTCAGC 539
 QY 1010 TCCCGCGCCAAAAAATGGCTCTCTCTGGGCTCCAGGATTTGTCTCCCACTCCAGC 1069
 D 538 TCCCGCGCCAAAAAATGGCTCTCTCTGGGCTCCAGGATTTGTCTCCCACTCCAGC 479
 QY 1070 ATCCCAAACTGGTACTCCCTGACCCAGGCGGCGCTTAAGCCAGGCGCTCTTGACAGGCTCCA 1129
 D 478 ATCCCAAACTGGNACTCCCTGACCCAGGCGGCGCTTAAGCCAGGCGCTCTTGACAGGCTCCA 420
 QY 1130 TGAGTCAAGCGCCCTCCCACTCCAGGACTCCAGATTTACCTCTCCCTGCGAGTCTGG 1189
 D 419 TGAGTCAAGCGCCCTCCCACTCCAGGACTCCAGATTTACCTCTCCCTGCGAGTCTGG 360
 QY 1190 GTTCCAGACTGAGTCTCTCCCAATCAGGCGCTTAGAGCCGCGCTCCCAACCTGG 1249
 D 359 GTTCCAGACTGAGTCTCTCCCAATCAGGCGCTTAGAGCCGCGCTCCCAACCTGG 300
 QY 1250 ACTCTGGACTTAGGCGCCCTTAAATCTAGACTTTCTCTTTTAGGTTTCAGGTTCTCCCTAT 1309
 D 299 ACTCTGGACTTAGGCGCCCTTAAATCTAGACTTTCTCTTTTAGGTTTCAGGTTCTCCCTAT 240
 QY 1310 GGTGCTCGGAAGTCTCTGAAAGTGGACTTTCTCAGGCTTGACCTGCCCGCCACCTATC 1369
 D 239 GGTGCTCGGAAGTCTCTGAAAGTGGACTTTCTCAGGCTTGACCTGCCCGCCACCTATC 180
 QY 1370 CCCGCGTGGAGCTGTGGGGCAGCAGATCAGAGCCACTGATAGGGCGCCCTAGAGGT 1429
 D 179 CCCGCGTGGAGCTGTGGGGCAGCAGATCAGAGCCACTGATAGGGCGCCCTAGAGGT 120
 QY 1430 ACAGGGTCTGCCAGCAGGTCGCCACGAGTGTCTCTCATTTTATTTTACGCTCCATTT 1489
 D 119 ACAGGGTCTGCCAGCAGGTCGCCACGAGTGTCTCTCATTTTATTTTACGCTCCATTT 60
 QY 1490 TGCCCATAGATGGCAGAGGGGTGAGATGGCTCATCCCTCCAGATTTCTGCAATAAA 1549
 D 59 TGCCCATAGATGGCAGAGGGGTGAGATTT-GCTCATCCCCCTTNNNDWHTTCTCTHWTWAA 1

JOURNAL
COMMENT

Unpublished
On Feb 13, 2001 this sequence version replaced gi:12792042.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 888.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD001BC02NP1&cluster=888.f. Contact :
Feng Liang Email : fliang@lifetech.com/ Invitrogen Corporation 1600
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DD001BC02NP1.

FEATURES

source

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD001YB04"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 253 a 301 c 362 g 249 t 36 others
ORIGIN

Query Match 47.8%; Score 942.6; DB 9; Length 1201;
Best Local Similarity 96.4%; Pred. No. 4.3e-229;
Matches 1008; Conservative 16; Mismatches 16; Indels 6; Gaps 6;

```

QY 475 GCCTGTGACACGACCTGGCTCCCGTTCTTTTGTGCGCGACAAAGGGTCTCACTACTGC 534
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1071 KGGTGTGACAGCCACTGGSTYCCCG-TCTTTTGTGCGCGAC-ARGGTGCTYAACTMTGC 1014
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 535 GTGCCCTCTATGAGAACAAAGTTTCTCTCGCTGCGCGCGGTGTCGACGAAGACGCTGACA 594
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1013 GTGCCCTCTATGAGAACAAAGTAACWCCTY-CGCGCGCGCTGTCGACGAAGACGCTGACA 955
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 595 CAGGTGTGAGTACATACCGTGATCAGCGTGGCATCGAGATGCTGGTGTGTACCGGA 654
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 954 CAGGTGTGAGTACATACCGTGATCAGCGTGGCATCGAGATGCTGGTGTGTACCGGA 895
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 655 TGCCAGACCCCTGCGCAGGCGAGCTTACCTCCCGGATGAAGATCCCTACTGTGTG 714
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 894 TGCCAGACCCCTGCGCAGGCGAGCTTACCTCCCGGATGAAGATCCCTACTGTGTG 835
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 715 GCCTGTTTGGAGAACTCTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATCGTAGGA 774
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 834 GCCTGTTTGGAGAACTCTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATCGTAGGA 775
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 775 CTCGGTGGAGGCAAGTATGTCTTTTGAAGACCCAGACTGGCACCACCACTGCTTCC 834
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 774 CTCGGTGGAGGCAAGTATGTCTTTTGAAGACCCAGACTGGCACCACCACTGCTTCC 715
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 835 TCGCCCGCTGCTCTACCTCCTCTGTGGCGCAGGCTTCGTACCGGATGGAGACCAAGT 894
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 714 TCGCCCGCTGCTCTACCTCCTCTGTGGCGCAGGCTTCGTACCGGATGGAGACCAAGT 655
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 895 CTCCTGCCAGGCTGTGAGCCAGGCGCCCTTAAGCCAGGCTCTTGGACCCAGGCTTTCC 954
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 654 CTCCTGCCAGGCTGTGAGCCAGGCGCCCTTAAGCCAGGCTCTTGGACCCAGGCTTTCC 595
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 955 CATACCAGGCGCCAGGACTGTGGCTCTTTTCTAAACCACTCTGGGACTCAGTCCCC 1014
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 594 CATACCAGGCGCCAGGACTGTGGCTCTTTTCTAAACCACTCTGGGACTCAGTCCCC 535
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1015 CGCCAAAAAATGGGTCTCTCTCTGGGCTCCAGGATGTCTCCCGCACTCCAGCATCCC 1074
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 534 CGCCAAAAAATGGGTCTCTCTCTGGGCTCCAGGATGTCTCCCGCACTCCAGCATCCC 475
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

QY 1075 CAAATGGTACTCCTCTGACCCAGGCCCCCAATCCTGTGGGTCTTACAGAGCCCTCCATGAGT 1134
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 474 CAAATGGTACTCCTCTGACCCAGGCCCCCAATCCT-GGCTCTTACAGAGCCCTCCATGAGT 416
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1135 CAAGCCCCCTCCACACACTGACCTGCAGACTCCAGAAATCACCTCTCCCTGCGAGTCTGGTTC 1194
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 415 CAAGCCCCCTCCACACAMMRAACTCCAGAAATTCACCTCTCCCTGCGAGTCTGGTTC 356
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1195 CAGACTGAGTCTCTCTCCCAATCAGGGCTCTAGACCCGAGCCCTCCAAACCTGACACTCT 1254
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 355 CAGACTGAGTCTCTCTCCCAATCAGGGCTCTAGACCCGAGCCCTCCAAACCTGACACTCT 296
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1255 GGGACTTGGCCCCCTTAAATCTAGACTTCTCTTATAGGTTTCAGGTTCTCTATGGGTG 1314
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 295 GGGACTTGGCCCCCTTAAATCTAGACTTCTCTTATAGTTCAGGTTCTCTATGGGTG 236
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1315 CTTGGGAGTCTCTTGAAGTGGACTGTCTCAGGCTTGACCTGCCGCCACCCCATCCCGC 1374
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 235 CTTGGGAGTCTCTTGAAGTGGACTGTCTCAGGCTTGACCTGCCGCCACCCCATCCCGC 176
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1375 GGTTCAGGCTGTGGGGGAGCAGATCAGGAGCCCACTGATAAGGGGCCCTAGGTAC-AG 1433
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 175 GGTTCAGGCTGTGGGGGAGCAGATCAGGAGCCCACTGATAAGGGGCCCTAGGTACAG 116
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1434 GGTCTGCCAGCAGCTGCCACCGAGTGT-CTTCTCAFTTTTATTTCAGCTCCATTTTC 1492
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 115 GGTCTGCCAGCAGCTGCCACCGAGTGTACATCTCATTTATTTCNCCTCCATTTTCG 56
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1493 CCATAGATGGGAGAGGGGTGAGATT 1518
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 55 CCATAGATGGGAGAGGGGTGAGATT 30
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 7

AL525660/c

LOCUS

DEFINITION

AL525660 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

AL525660

AL525660.2 GI:31063524

EST.

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li W.B., Gruber C., Jessee J. and Polayes D.

Full-length cDNA libraries and normalization

Unpublished

On Feb 13, 2001 this sequence version replaced gi:12789153.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 888.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DD0012CA04NP1&cluster=888.f. Contact :

Feng Liang Email : fliang@lifetech.com/ Invitrogen Corporation 1600

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DD0012CA04NP1.

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DD001YB07"

/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

FEATURES

source

QY 1309 TGGTGCCTGGGAAGTCCCTTGAAGTGGAGCTGTCTCAGGCTTGAGCTGCCTGCCACACCCCAT 1368
Db 232 TGGTGCCTGGGAAGTCCCTTGAAGTGGAGCTGTCTCAGGCTTGAGCTGCCTGCCACACCCCAT 173
QY 1369 CCCCGCGGTGAGGCTGTGGGGCAGCAGATCAGGAGCCACTGATGAAGGGGCCCTTAGG 1428
Db 172 CCCCGCGGTGAGGCTGTGGGGCAGCAGATCAGGAGCCACTGATGAAGGGGCCCTTAGG 113
QY 1429 TACAGGGTGTGCCAGCAGGTCGCCACCGAGTGTCTTCTCATTTTATTTTACGCTCCATT 1488
Db 112 TACAGGGTGTGCCAGCAGGTCGCCACCGAGTGTCTTCTCATTTTATTTTACGCTCCATT 53
QY 1489 TTGCCCATAGATGGGAGGGGTGAGATGGCTCATCCCTTCACGATTTCT 1541
Db 52 TTGCCCATAGATGGGAGGGGTGGA---TTGCTCATCCCTTBNNYCTCT 3

RESULT 10

AL526733

LOCUS

DEFINITION AL526733 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC003YC05 5-PRIME, mRNA sequence.

ACCESSION

AL526733

VERSION

AL526733.2

GI:31041994

EST.

KEYWORDS

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

On Feb 13, 2001 this sequence version replaced gi:12787226.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 888.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/ccluster.cgi?seq=CS0DC003AB03QP1&cluster=888.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DC003AB03QP1.

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DC003YC05"

/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 239 a 364 c 311 g 248 t 39 others

ORIGIN

Query Match

Best Local Similarity 92.2%; Pred. No. 5.3e-219;

Matches 989; Conservative 24; Mismatches 49; Indels 11; Gaps 6;

QY 1 GGTCCCTGTCAGTCGGCAGCTCGCCGAGCTCGCTCGGCCCGCGCTTGGCCGCCGCC 60

Db 70 GTTCCCTGTCAGTCGGCAGCTCGCCGAGCTCGCTCGGCCCGCGCTTGGCCGCCGCC 129

QY 61 GCAGGGTTCTCTCCCTTCCGACCATGACGAGTCAATTTGACTGTGCAAAATGCAACGAG 120

Db 130 GCA--GTTCTCTCCCTTCCGACCATGACGAGTCAATTTGACTGTGCAAAATGCAACGAG 187

QY 121 TCCTGTATGGACGCAAGTACATCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTAT 180
Db 188 TCCTGTATGGACGCAAGTACATCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTAT 247
QY 181 GACAATACCTTTGCCAACACCTGTGTAGTGCAGCAGCTTATCGGGCATGACTCGAG 240
Db 248 GACAATACCTTTGCCAACACCTGTGTAGTGCAGCAGCTTATCGGGCATGACTCGAG 307
QY 241 GAGCTGTCTTATGAAGACCGCCATTTCCACAGGGGTGCTTCCGCTGCTCCGCTGCCAG 300
Db 308 GAGCTGTCTTATGAAGACCGCCATTTCCACAGGGGTGCTTCCGCTGCTCCGCTGCCAG 367
QY 301 CGCTCACTAGCCGATGAACCTTCACTGCGCAGACAGTGTGTCTCTCAATGACTGC 360
Db 368 CGCTCACTAGCCGATGAACCTTCACTGCGCAGACAGTGTGTCTCTCAATGACTGC 427
QY 361 TACTGCAGTGGCTTTTCTCGCAGTGTCTCGCTGTGGGAGAGTGTCTCATGCTGGTCC 420
Db 428 WACTGCAGTGGCTTTTCTCGCAGTGTCTCGCTGTGGGAGAGTGTCTCATGCTGGTCC 487
QY 421 CGGAAGCTGGAATATGAGGCCAGACATGGCAGTGTGTCTCTCTGCTGAGTGGCTGT 480
Db 488 CGGAAGCTGGAATATGAGGCCAGACATGGCAGTGTGTCTCTCTGCTGAGTGGCTGT 547
QY 481 GAACGCCACTGGGCTCCCGTCTTTTGTGCGCGACAAGGGTGTCTCACTACTGCGTCCC 540
Db 548 GAACGCCA-TGGGCTCCCGTCTTTTGTGCGCGACAAGGGTGTCTCACTACTGCGTCCC 606
QY 541 TGTATGAGAACAAAGTTTGTCTCGTGGCGCCGCTGCGAGAGAGCTGTGACACAGGTT 600
Db 607 TGTATGAGAACAAAGTTTGTCTCGTGGCGCCGCTGCGAGAGAGCTGTGACACAGGTT 666
QY 601 GGAGTGACATACCGTGTATGAGCGTGTGAGTGTGTCTGTGTGTGTGTGTGTGTGTGTGT 660
Db 667 GGAGTGACATACCGTGTATGAGCGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 726
QY 661 ACSCCCTTGGCAGGCGAGCTTCACTCCCGGAGTGAAGTCACTGTGTGTGTGTGTGTGTGT 720
Db 727 AMSCCCTTGGCAGGCGAGCTTCACTCCCGGAGTGAAGTCACTGTGTGTGTGTGTGTGTGT 786
QY 721 TTTGGAGAACTCTTTGACACCTAAGTGTGAGCAGCTGTGAAGCGCCGCTTGTGTGTGTGTGT 780
Db 787 TTTGGAGAACTCTTTGACACCTAAGTGTGAGCAGCTGTGAAGCGCCGCTTGTGTGTGTGTGT 846
QY 781 GGAGGCAACTATCTGTCTTGAAGACCGAGCTGTGAGCAGCTGTGAGCAGCTGTGAGCAGCT 840
Db 847 GGAGGCAAGTATGTGTCTTTGAAGACCGAGCTGTGAGCAGCTGTGAGCAGCTGTGAGCAGCT 906
QY 841 CGCTGTCTACCTCCCTGTGTGGCGCAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 907 CGCTGTCTACCTCCCTGTGTGGCGCAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 966
QY 901 CAGGGGTGTAGCAGGCGAGGCGCTTAAAGCAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Db 967 CAGGGGTGTAG-CAGGCGAGGCGCTTAAAGCAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
QY 961 ACGGGCGCAGGCTGTGGCTTCTTCTTAAACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db 1023 MACGGCGCAGGCTGTGGCTTCTTCTTAAACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1078
QY 1021 AAAAAAATGGGTCTCTTCTTGGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1073
Db 1079 AATKGGTCTTCTTGGSCCAGRATTTTTCMCTCASCAYCCMAACTGKMCYCY 1131

RESULT 11

AL526734/c

LOCUS

DEFINITION

AL526734 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

cDNA clone CS0DC020YH03 3-PRIME, mRNA sequence.

ACCESSION

AL526734

VERSION

AL526734.2

GI:31064592

KEYWORDS

EST.

AL526734 1200 bp mRNA linear EST 23-MAY-2003
AL526734 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC020YH03 3-PRIME, mRNA sequence.
AL526734
AL526734.2 GI:31064592
EST.

primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT	266 a	286 c	371 g	218 t	60 others
ORIGIN					
Query Match	45.4%; Score 894.6; DB 13; Length 1201;				
Best Local Similarity	92.7%; Pred. No. 7.3e-217;				
Matches 1002;	Conservative 37; Mismatches 31; Indels 11; Gaps 10;				
Qy	437	GAGGCCAGACATGGCATGACGACCTGCTCTCTGTGTCAGTGGCTGTGAACAGCCACTGGGCT	496		
Db	1096	GRGSCARASMTTGGATTTRASACTTSTCTCTGTGTSAGTG--TGTGACGCCACTGGGCT	1039		
Qy	497	CCGCTCTCTTTGTGCCCGCAAGGGTCTCACTACTCGCTGCTGTATGAGAACAGT	556		
Db	1038	CCSKYITTTTGTCCCMACACGGGTGCTCASCMTCTGSKGCCCTGTATRGGAARAKT	979		
Qy	557	TTCGTCCTCGTGGCCCGCTGTCAGCAAGCGCT--GACACAGGTGGAGTGACATACCGT	615		
Db	978	TTCGTCCT--TCGCGCCGCTGTCAGCAAGCGSTKGACACAGGKGAGTGACATACCGT	920		
Qy	616	-GATCAGCCGTGGCATCG--AGAATGCTGTGTCTGTAC--GGATGTCAGACGCCCC--TGGC	671		
Db	919	GGATCAGCCGTGGCATCGRGAATGTCTGTGTAGACCGGGATGCCAKAGCCCKCTGGC	860		
Qy	672	AGGCGACAGTTCACCTCCCGGATGAAGATCCCTACTGCTGTGCTGTGGAGAA--C	730		
Db	859	AGGCGACAGTTCACCTCCCGGATGAAGATCCCTACTGCTGTGCTGTGGAGAA--C	800		
Qy	731	TCCTTTCACCTTAAGTGCAGCAGCTGCAAGGCCGCCA--TCGTAGGACTCGGTGGAGGCAAG	789		
Db	799	TCCTTTCACCTTAAGTGCAGCAGCTGCAAGGCCGCCA--TCGTAGGACTCGGTGGAGGCAAG	740		
Qy	790	TATGTGTCCTTTGAAGACGACACTGGCAGACACAACTGCTTCCTCGCCCGCTGTCT	849		
Db	739	TATGTGTCCTTTGAAGACGACACTGGCAGACACAACTGCTTCCTCGCCCGCTGTCT	680		
Qy	850	ACCTCCCTGTGGGCCAGGCTTCGTACCGGATGGAGACCAAGTCTGCCAGGGCTGT	909		
Db	679	ACCTCCCTGTGGGCCAGGCTTCGTACCGGATGGAGACCAAGTCTGCCAGGGCTGT	620		
Qy	910	AGCCAGCAGGGCCCTTAAGCCAGGCTCTTGAGCCGCCAGGCTTTCCCATACACAGGGCCCA	969		
Db	619	AGCCAGCAGGGCCCTTAAGCCAGGCTCTTGAGCCGCCAGGCTTTCCCATACACAGGGCCCA	560		
Qy	970	GGACTGTGGCTCTCTTTTCTAAACACCTCTGGGACTCAGCTCCCGCCGCAAAAATG	1029		
Db	559	GGACTGTGGCTCTCTTTTCTAAACACCTCTGGGACTCAGCTCCCGCCGCAAAAATG	500		
Qy	1030	GGTCTCTCTGGGCTCCAGGATTTCTCCCACTCCAGATCCCAAACTGGTACTCC	1089		
Db	499	GGTCTCTCTGGGCTCCAGGATTTCTCCCACTCCAGATCCCAAACTGGTACTCC	440		
Qy	1090	TGACCCAGGCCCCCAATCTCTGGCTTTTACAGAGCCTTCCATGAGTCAAGCCCCCTCCCA	1149		
Db	439	TGACCCAGGCCCCCAATCTCTGGCTTTTACAGAGCCTTCCATGAGTCAAGCCCCCTCCCA	381		
Qy	1150	CACCTGACATCCAGATTCACCTCTCCCTCTGAGTCTGGTTCCTCCAGACTGAGTCTCT	1209		
Db	380	CACCTGACATCCAGATTCACCTCTCCCTCTGAGTCTGGTTCCTCCAGACTGAGTCTCT	321		
Qy	1210	CCCCAAATCAGGGCTCTAGACCCGAGCCCTCCAAACCTGGACTCTGGGACTTGGCCCCC	1269		
Db	320	CCCCAAATCAGGGCTCTAGACCCGAGCCCTCCAAACCTGGACTCTGGGACTTGGCCCCC	261		
Qy	1270	TAAATCTAGACTTCTCTTTATAGGTTTCAGGTTCTCTATGGGTGCTTGGGAAGTCTTG	1329		
Db	260	TAAATCTAGACTTCTCTTTATAGGTTTCAGGTTCTCTATGGGTGCTTGGGAAGTCTTG	201		
Qy	1330	AAAGTGACATGTTCTCAGGCTTGACCTGCCGCCACCCCATCCCGCGGTTGAGGCTG	1389		
Db	200	AAAGTGACATGTTCTCAGGCTTGACCTGCCGCCACCCCATCCCGCGGTTGAGGCTG	141		

Qy	1390	GGCAGCAGATCAGGAGCCCACTGATATAGGGGCCCTAGGTTACAGGTGCTGCCACGAGG	1449	
Db	140	GGCAGCAGATCAGGAGCCCACTGATATAGGGGCCCTAGGTTACAGGTGCTGCCACGAGG	81	
Qy	1450	TCGCCACCGAGTGCTCTCATTTTATTTTCAGCTCCATTTTGCCCATAGATGGCAGAGG	1509	
Db	80	TCCKCWCAGAGTGCTCTCATTTTATTTTCAGCTCCATTTTGCCCATAGATGGTCA	21	
Qy	1510	G 1510		
Db	20	G 20		
RESULT 13				
AL551558/c				
LOCUS				
DEFINITION AL551558 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA				
clone CSODI063YD23 3-PRIME, mRNA sequence.				
ACCESSION AL551558				
VERSION AL551558.2 GI:31273374				
KEYWORDS EST.				
SOURCE Homo sapiens (human)				
ORGANISM Homo sapiens				
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.				
TITLE Full-length cDNA libraries and normalization				
JOURNAL Unpublished				
COMMENT On Feb 15, 2001 this sequence version replaced gi:12889621.				
Contact: Genoscope - Centre National de Sequencage				
BP 191 91006 EVRY cedex - France				
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr				
Library was constructed by Life Technologies, a division of				
Invitrogen. This sequence belongs to sequence cluster 888.f For				
more information about this cluster, see				
http://www.genoscope.cns.fr/				
cgi-bin/cluster.cgi?seq=CSODI063CB12NP1&cluster=888.f. Contact :				
Feng Liang Email : fliang@lifetech.com URL :				
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600				
Faraday Avenue Genoscope sequence ID : CSODI063CB12NP1.				
Location/Qualifiers				
1..1116				
/organism="Homo sapiens"				
/mol_type="mrna"				
/db_xref="taxon:9606"				
/clone="CSODI063YD23"				
/tissue_type="PLACENTA COT 25-NORMALIZED"				
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"				
/note="1st strand cDNA was primed with a NotI-oligo(dT)				
primer. Five prime end enriched, double-strand cDNA was				
digested with Not I and cloned into the Not I and EcoR V				
sites of the pCMVSPORT 6 vector. Library was normalized."				
BASE COUNT 235 a 259 c 337 g 226 t 59 others				
ORIGIN				
Query Match 45.4%; Score 894.2; DB 9; Length 1116;				
Best Local Similarity 94.0%; Pred. No. 9.2e-217;				
Matches 913; Conservative 26; Mismatches 27; Indels 5; Gaps 2;				
Qy	548	AGACACAGTTTGCTCTCGTCCGCTGCAGACAGCCGCTGCAGACAGGTTGAGTCA	607	
Db	984	AGAGAAAAAGTTGTSTCCRYCGSCCGTTSMAGMAAGACGCTGCACACAGGGTGGAKTGA	925	
Qy	608	CATACCGTGTATCAGCCGCTGGCATCGAGAATGTCTGCTGTCTACCGGATGCCAGACGCC	667	
Db	924	CATACCGTGTATCAGCCATGTCATCGAGAATGTCTGCTGTCTACCGGATGCCAGACGCC	865	
Qy	668	TGGCAGGCGCAGCTTCACTCCCGGGATGAAGATCCCTACTGTGTGGCTGTGTTGGAG	727	
Db	864	TGGCAGGCGCAGCTTCACTCCCGGGATGAAGATCCCTACTGTGTGGCTGTGTTGGAG	805	

QY	728	AACTCTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATCTGTAAGGACTCGGTGGAGGCA	787
Db			
Db	804	AACTCTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATC----	749
QY	788	AGTATGTGTCCTTTGAAGACCCACACTGCGACACCAACTGCTTCTCTCGGCCCGCTGCT	847
Db	748	AGTATGTGTCCTTTGAAGACCCACACTGCGACACCAACTGCTTCTCTCGGCCCGCTGCT	689
QY	848	CTACTCTCCCTGTGGGCCAGGGCTTCGTACCGGATGGAGACCAAGTGTCTGTGCCAGGGCT	907
Db	688	CTACTCTCCCTGTGGGCCAGGGCTTCGTACCGGATGGAGACCAAGTGTCTGTGCCAGGGCT	629
QY	908	GTAGCCAGCAGGGCCCTAAGCCAGGGCTCTCTGGACCCAGGCTTTCCCATACACACGGGC	967
Db	628	GTAGCCAGCAGGGCCCTAAGCCAGGGCTCTCTGGACCCAGGCTTTCCCATACACACGGGC	569
QY	968	CAGGAGCTGTGGCTCTCTTTCTAAACCACTCTCTGGGACTCAGTCTCCCGCCCAAAAAA	1027
Db	568	CAGGAGCTGTGGCTCTCTTTCTAAACCACTCTCTGGGACTCAGTCTCCCGCCCAAAAAA	509
QY	1028	TGGGTCTCTCTTGGGCTCCAGGATGTCTCCCGACTCCAGCATCCCAAACTGGTACTC	1087
Db	508	TGGGTCTCTCTTGGGCTCCAGGATGTCTCCCGACTCCAGCATCCCAAACTGGTACTC	449
QY	1088	CCTGACCCAGGCCCCCAATCTCTGGGCTCTTACAGAGCCTTCCATGATCAAGCCCTCC	1147
Db	448	CCTGACCCAGGCCCCCAATCTCTGGGCTCTTACAGAGCCTTCCATGATCAAGCCCTCC	390
QY	1148	CACACTGGACTCCAGATTCACCTCTCCCTCTGAGTCTGGGTTCCAGACTGATCTCT	1207
Db	389	MACACTGKACTCCAGATTCACCTCTCCCTCTGAGTCTGGGTTCCAGACTGATCTCT	330
QY	1208	CTCCCAATCAGGGCTCTAGACCCGAGCCCTCCAACTCTGGACTCTGGAGCTTAGGCC	1267
Db	329	CTCCCAATCAGGGCTCTAGACCCGAGCCCTCCAACTCTGGACTCTGGAGCTTAGGCC	270
QY	1268	CCTTAAATCAGACTTCTCTTTATAGTTTTCAGGCTCTCTATGGTGTCTGGGAGTCTCT	1327
Db	269	CCTTAAATCAGACTTCTCTTTATAGTTTTCAGTCTCTATGGGTTCCTGKGAATCTCT	210
QY	1328	TGAAGTGGACTGTCTCAGGCTTGACCTGCCCCCAGCCCATCCCGCGGTTGAGGCTGTG	1387
Db	209	TGAAGTGGACTGTCTCAGGCTTGACCTGCCCCCAGCCCATCCCGCGGTTGAGGCTGTG	150
QY	1388	GGGGCAGCAGATCAGGAGCCCACTGATAAGGGGCCCTAGGGTACAGGGTGTCTGCCAGCA	1447
Db	149	GGGGCAGCAGATCAGGAGCCCACTGATAAGGGGCCCTAGGGTACAGGGTGTCTGCCAGCA	90
QY	1448	GGTCCGCCAGGAGTGTCTCTCATTTTATTTTCAGCTTCCATTTTGCCCATAGATGGCAGA	1507
Db	89	GGTCCGCCAGGAGTGTCTCTCATTTTATTTTCAGTCTCCATTTTGCCCATAGATGRACAGA	30
QY	1508	GGGGTGCAGATT 1518	
Db	29	NANNTNNATTT 19	
RESULT 14			
BX445770/c			
LOCUS	BX445770	1201 bp	mRNA linear EST 15-MAY-2003
DEFINITION	clone CS0DH006Yall 3-PRIME, mRNA sequence.		
ACCESSION	BX445770		
VERSION	BX445770.1 GI:30774361		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		

Unpublished	
Contact: Genoscope	
Genoscope - Centre National de Sequencage	
BP 191 91006 EVRY cedex - France	
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr	
Library was constructed by Life Technologies, a division of	
Invitrogen. This sequence belongs to sequence cluster 888.f For	
more information about this cluster, see	
http://www.genoscope.cns.fr/	
cqi-bin/cluster.cgi?seq=CS1DH002ZC03NP1&cluster=888.f. Contact :	
Feng Liang Email : fliang@lifetech.com URL :	
http://fulllength.invitrogen.com/Invitrogen Corporation 1600	
Faraday Avenue Genoscope sequence ID : CS1DH002ZC03NP1.	
Location/Qualifiers	
1. 1201	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="CS0DH006Yall"	
/tissue_type="T CELLS (JURKAT CELL LINE)"	
/cell_line="JURKAT CELL LINE"	
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"	
/note="Vector: pCMVSPORT.6; 1st strand cDNA was primed	
with a NotI-oligo(dT) primer. Five prime end enriched, into	
double-strand cDNA was digested with Not I and cloned, into	
the Not I and EcoRV sites of the pCMVSPORT 6 vector.	
Library was not normalized."	
BASE COUNT	259 a 293 c 349 g 229 t 71 others
ORIGIN	
Query Match	45.4%; Score 894.2; DB 13; Length 1201;
Best Local Similarity	90.6%; Pred. No. 9.3e-217;
Matches	969; Conservative 32; Mismatches 64; Indels 5; Gaps 5;
QY	461 GCTTCTGTCAGTGGCTGTGAACAGCCACTGGGCTCCCGTCTTTTGTGCGCGACAAG 520
Db	1068 KFTTCTTGGGKADGKTTKAAAVCAAHGGTBCCGTGTGTTTSCCCAAAGG 1009
QY	521 GTGCTCACTACTGCGCTGCTATGAGAACAAAGTTGCTGCTGCGCGCGCTGCA 580
Db	1008 GGGTTAAACATGCKKCCCGGTTTARGAACARTTTKCCCTTCGSSCGSGSCR 949
QY	581 GCAAGACGCTACACAGGTTGAGTACATACCGTATCAGCGTGGCATCGAGATGTC 640
Db	948 GCAAAAGGTTAAACAGGGGGAATTAACATACCGTGTMMAGCCGTGGCATCGARAATTC 889
QY	641 TGCTCTGACCGGATGCCAGACGCCCTTGGCAGGCGAGTTCACCTCCCGGATGAG 700
Db	888 T-GTTTKTACCGGATGCCAACGCCCTTGGCAGGCGAGTTCCTCCCTCCGGATGAG 830
QY	701 ATCCCTACTGTGTGCGCTGTTTGGAGAACTCTTTGCACCTAAGTGCAGCAGCTGCAAG 760
Db	829 ATCCCWACTTTTGGCTGTTTGGAGAACTCTTTGCACCTAAGKGDGCGCAGCTGCAAG 770
QY	761 GCCCCATCGTAGGACTCGGTGGAGGCAAGTATGTGCTTTTGAAGACCGACACTGGCACC 820
Db	769 GCCCCATCGTAGAACTCGGTGRAGGCAAGTTTGTGCTTTTGAAGACCGAWACTGGCACT 710
QY	821 ACAACTGCTTCTCTGCGGCCCGCTGCTACCTCCCTGCTGGCGCAGGGCTTCGTACCGG 880
Db	709 ACAACTGCTTCTCTGCGGCCCGCTGCTACCTCCCTGCTGGCGCAGGGCTTCGTACCGG 650
QY	881 ATGAGAGCAAGTGTCTGCCAGGGCTGTAGCAGGAGGCGCTTAAAGCAGGGCTTCCTG 940
Db	649 ATGGAGAGCAAGTGTCTGCCAGGGCTGTAGCAGGAGGCGCTTAAAGCAGGGCTTCCTG 590
QY	941 GACCCAGGCTTTCCCATACCCAGGCGCCAGGACTGTGGCTCTCTTTTCTAAACACACTCTG 1000
Db	589 GACCCAGGCTTTCCCATACCCAGGCGCCAGGACTGTGGCTCTCTTTTCTAAACACACTCTG 530
QY	1001 GGACTCAGCTCCCCCGCC-AAAAAATGGGTCTCTCTCTGGGCTCCAGGATTTGCTCC 1059
Db	529 GAATCTCAGCTCCCCCGCCAAAAAATGGGTCTCTCTCTGGGCTCCAGGATTTGCTCC 470

Db ||| | ||||| | | |||||
 984. CCGGCCAAAAAATGGGCCTTCT 1009

Search completed: August 27, 2003, 14:11:44
Job time : 2843 secs